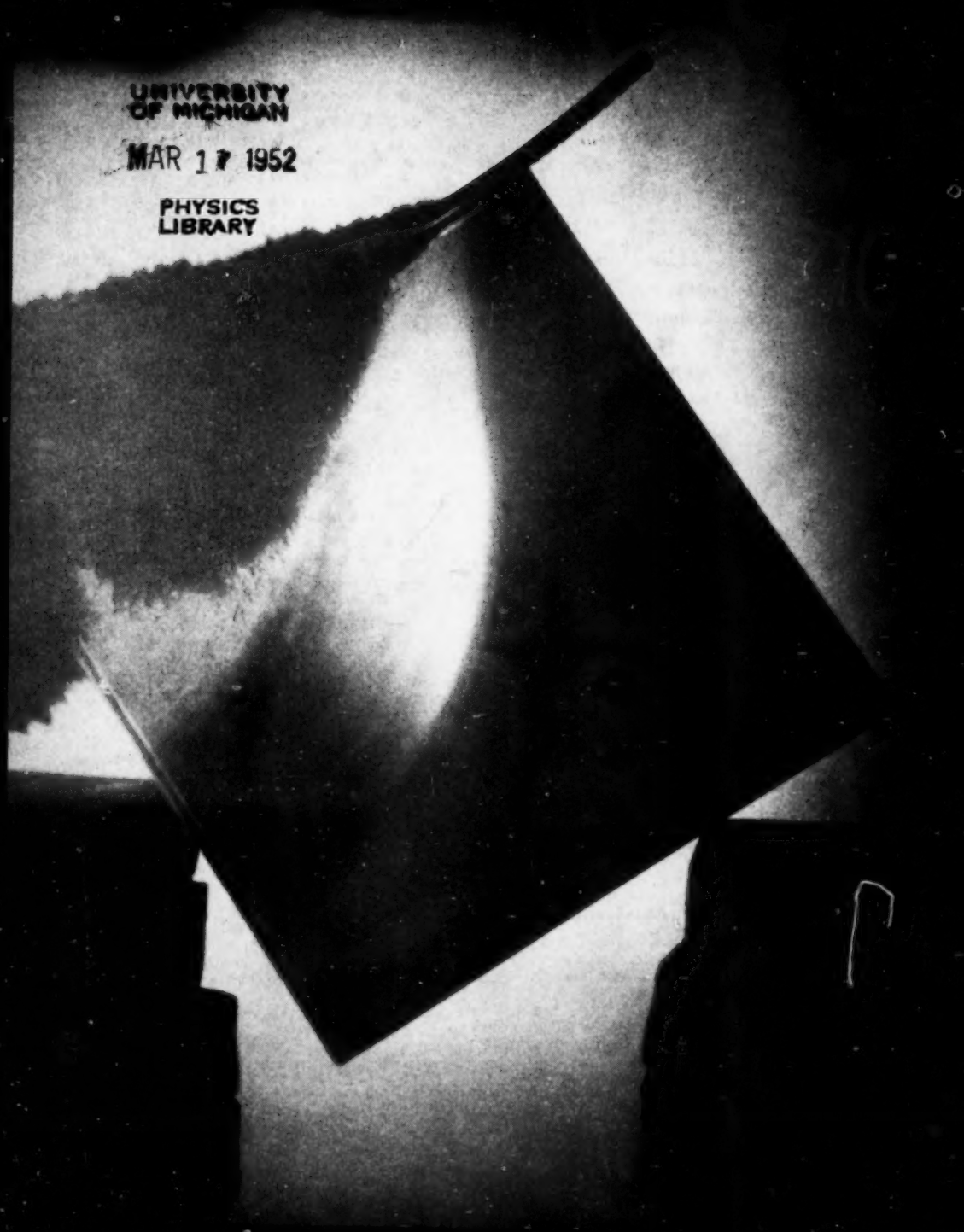
JOURNAL OF APPLIED PHYSICS



UNIVERSITY OF MICHIGAN
MAR 17 1952
PHYSICS LIBRARY

JOURNAL OF APPLIED PHYSICS

Vol. 23, No. 1 JANUARY, 1952

In This Issue

	Page
Vertical Shaft for the Production of Thick Artificial Clouds and for the Study of Precipitation Mechanics — Ross Gunn	1
Mass-Spectrographic Study of Ionic Reactions in Hydrogen — Raymond L. Murray	6
Experimental Investigation of the Dielectric Rod Antenna of Circular Cross Section Excited in Rotationally Symmetrical Modes — C. M. McKinney	11
Collapse and Rebound of a Gas Bubble — Leon Trilling	14
On the Empirical Law of Adhesive Wear — J. T. Burwell and C. D. Strang	18
Electromagnetic Resonant Behavior of a Confocal Spheroidal Cavity System in the Microwave Region — J. C. Simons and J. C. Slater	29
Structure of Sputtered Silver Films — C. E. Ells and G. D. Scott	31
Simulation of Trajectories of Charged Particles in Magnetic Fields — George H. Vineyard	35
Supersaturation of Nitrogen in a Hypersonic Wind Tunnel — I. Faro, T. R. Small, and F. K. Hill	40
Investigation of Stress-Strain Relations of Metal Wires by Electrical Resistance Changes — Irwin Vigness	43
Glass Cracking Caused by High Explosives — E. M. Pugh, R. v. Heine-Geldern, S. Foner, and E. C. Mutschler	48
M. I. T. Linear Electron Accelerator — P. T. Demos, A. F. Kip, and J. C. Slater	53
Particle Dynamics in the Linear Accelerator — J. R. Terrall and J. C. Slater	66
Field Strength Measurements in Resonant Cavities — L. C. Maier, Jr., and J. C. Slater	68
Determination of Field Strength in a Linear Accelerator Cavity — L. C. Maier, Jr., and J. C. Slater	78
Forced Thickness-Shear and Flexural Vibrations of Piezoelectric Crystal Plates — Raymond D. Mindlin	83
Thermal Dependence of the Elastic Moduli of Polycrystalline Graphite — F. E. Faris, L. Green, Jr., and C. A. Smith	89
Nonsteady Heat Diffusion Problem with Spherical Symmetry — M. S. Plesset and S. A. Zwick	95
Temperature Dependence of Viscosity for Water and Mercury — Preston M. Kampmeyer	99
Instantaneous Power Spectra — Chester H. Page	103
Spectral Emissivity of Iron-Nickel Alloys — H. B. Wahlin, Robert Zentner, and James Martin	107
Thermal and Shot Fluctuations in Electrical Conductors and Vacuum Tubes — Salim S. Solomon	109
Heat Transfer in Isotropic Turbulence — Stanley Corrsin	113
X-Ray Measurement of Order in CuPt — C. B. Walker	118
Note on Critical Reflections of Elastic Waves at Free Surfaces — J. N. Goodier and R. E. D. Bishop	124
Direct-Current Network Analyzer for Solving Wave-Equation Boundary-Value Problems — George W. Swenson, Jr., and Thomas J. Higgins	126
Phenomena Associated with the Flight of Ultra-Speed Pellets. Part I Ballistics — William A. Allen, John S. Rinehart, and W. C. White	132
Probability Distribution of the Phase of the Resultant Vector Sum of a Constant Vector Plus a Rayleigh Distributed Vector — Kenneth A. Norton, Edna L. Schultz, and Helen Yarbrough	137
Steady-State One-Dimensional Flow in Rocket Motors — E. W. Price	142
Direct Determination of the Flow Curves of Non-Newtonian Fluids — Irvin M. Krieger and Samuel H. Maron	147
Electron Diffraction Evidence for the Existence of Microstress in Evaporated Metal Films — Eber K. Halteman	150
Letters to the Editor:	
Size of Latex Particles by X-Ray Scattering — B. R. Leonard, Jr., J. W. Anderegg, Paul Kaesberg, and W. W. Beeman	152
Effect of Pressure on Scintillation Phosphors — A. J. Reinsch and H. G. Drickamer	152
Direct Printing of Shadowed Electron Micrographs — R. M. Fisher	153
Dielectric Constants of Various Gases at 9470 Mc — Clayton M. Zieman	154
Physical Theory of Rubber Reinforcement — A. M. Bueche	154
Announcements	156
Program of the Electron Microscope Society of America	156
Index to Advertisers	xxvi

BOARD OF EDITORS

Elmer Hutchisson, *Editor*

Earle C. Gregg, *Assistant Editor*

Associate Editors

Howard H. Aiken
Ray E. Bolz
Lyle D. Borst
Robert D. Heidenreich
James Hillier
Leonard B. Loeb
W. James Lyons
Eric Reissner
Roman Smoluchowski

Publication of the American Institute of Physics

OFFICERS

George R. Harrison, *Chairman*
George B. Pegram, *Treasurer*

GOVERNING BOARD

J. W. Beams
C. Paul Boner
Wallace R. Brode
W. F. Fair, Jr.
S. A. Goudsmit
G. P. Harnwell
George R. Harrison
F. V. Hunt
Rudolph Kingslake
Paul Kirkpatrick
F. W. Loomis
J. R. Oppenheimer
George B. Pegram
I. I. Rabi
Duane Roller
Harold K. Schilling
Frederick Seitz
Francis G. Slack
John C. Steinberg

Consultant on Printing

Melvin Loos

ADMINISTRATIVE STAFF

Henry A. Barton, *Director*
Wallace Waterfall, *Executive Secretary*
Theodore Vorburger, *Advertising Manager*
Ruth F. Bryans, *Publication Manager*
Betty H. Goodfriend, *Assistant Secretary*
Eileen B. Neuberger, *Circulation Manager*
Kathryn Setze *Assistant Treasurer*

The Journal of Applied Physics, published monthly at Prince and Lemon Sts., Lancaster, Pa., is devoted to physics in its role as the science basic to other natural sciences and to the arts and industries. Previous to 1937 this publication was known as Physics. The Journal of Applied Physics publishes editorials and reviews, as well as technical papers of applied physics. Articles appearing in it are indexed regularly, according to their field, by Physics Abstracts, Chemical Abstracts, Engineering Index, Inc., Industrial Arts Index, etc.

Manuscripts should be submitted to Elmer Hutchisson, Editor, Case Institute of Technology, Cleveland 6, Ohio. Unless otherwise stated, submission of a manuscript is a representation that it has been neither copyrighted, published, nor submitted for publication elsewhere. The authors' institutions are requested to pay a **publication charge** of $8 per page which, if honored, entitles them to 100 reprints without further charge. Instructions will be sent with galley proofs.

Subscriptions, renewals, and orders for back numbers should be addressed to the American Institute of Physics, 57 East 55 Street, New York 22, New York.

Subscription Price	U.S. and Canada	Elsewhere
To members of the American Institute of Physics and Affiliated Societies	$10.00	$11.00
To all others	12.00	13.00

Back Number Prices

Yearly back number rate when complete year is available: $10.00. Single copies: $1.00 each.

Changes of address should be sent to the American Institute of Physics; **corrected proofs** should be sent to the Publication Manager, American Institute of Physics. **Advertising rates** supplied on request. Orders, advertising copy, and cuts should be sent to the American Institute of Physics.

Entered as second class matter January 22, 1937, at the Post Office at Lancaster, Pennsylvania, under the Act of March 3, 1879. Accepted for mailing at the special rate of postage provided for in paragraph (d-2), section 34.40, P. L. & R. of 1948, authorized May 2, 1932.

Copyright 1952, by the American Institute of Physics.

Journal
of
Applied Physics

Volume 23, Number 1 January, 1952

A Vertical Shaft for the Production of Thick Artificial Clouds and for the Study of Precipitation Mechanics

Ross Gunn
United States Weather Bureau, Washington, D. C.
(Received June 18, 1951)

Progress in cloud physics has been seriously limited because of the impossibility of conducting controlled cloud experiments on a sufficiently large scale. To meet the requirements, a laboratory has been built in which thick artificial expansion type clouds may be produced under controlled and measurable conditions.

A mine shaft 0.210 kilometer deep and 2.2×2.2 meters in cross section has been sealed and adapted to the production of thick artificial clouds. The shaft and its principal auxiliaries necessary for the study of some of the basic problems of clouds and precipitation are described. Dense, thick clouds completely filling the shaft are now being produced on a routine basis. Some preliminary illustrative data are presented.

INTRODUCTION

THE many problems that have emerged from present day studies of cloud physics and precipitation have reemphasized the necessity for developing new and better tools to advance such work. Thousands of cloud modification experiments in the free atmosphere have been made during the last few years and their interpretation attempted without even a reasonable knowledge of the physical behavior of the clouds or their past histories. Our own early experiments on the seeding of natural clouds emphasized the weaknesses of presently employed procedures and pointed to the necessity for further carefully controlled studies. Before much progress can be made on the problems of rain production from clouds at above freezing temperatures, two difficult physical problems must be solved in directly measurable experimental terms, namely, (a) In what way does the size and rate of growth of a raindrop depend upon the characteristics of the parent cloud, and (b) How much does the drop evaporate before it reaches the ground? Both of these problems require a fundamental understanding of basic mechanisms. It soon became evident to the author that in order to secure reliable, quantitative, experimental answers to these questions, the processes of rain production and evaporation had to be brought into the laboratory where conditions could be carefully controlled and direct physical measurements made. Moreover, new methods of measurement and new facilities had to be provided. It is the purpose of this paper to report the first success in producing thick artificial clouds and describe a new cloud laboratory built by the U. S. Weather Bureau to produce such thick expansion-type clouds under known and carefully controlled conditions. These clouds are currently being used to provide the data necessary to solve the first problem. A solution to the second problem has already been reached with the aid of the vertical shaft and new methods of measurement that will be presently described.[1–3]

BASIC REQUIREMENTS

In order to evaluate the detailed characteristics of clouds and determine in specific terms the mechanisms operating to produce a raindrop from a cloud, it is necessary that any artificial clouds employed for scientific study closely simulate the natural ones. The physical behavior of natural clouds is determined by their (a) age, (b) temperature, (c) water content, (d) the speed of their production, (e) the number and character of the nuclei, (f) the ionization of the ambient air, and (g) droplet size distribution. When a cloud is artificially produced by the expansion of air in a vessel, all of these characteristics except age can be easily controlled over a moderate range in the laboratory. It appears, therefore, that a great deal of basic information

[1] G. D. Kinzer and Ross Gunn, J. Meteor. 8, 71–83 (April, 1951).
[2] Ross Gunn and G. D. Kinzer, J. Meteor. 6, 243–248 (August, 1949).
[3] Ross Gunn, Rev. Sci. Instr. 20, 291–296 (1949).

might be learned from carefully conducted laboratory studies. In planning such studies, the experimenter appeared to be limited in two fundamental respects. The most serious limitation arises from the necessity for nearby walls. These walls increase the exchange of heat and radiation between the cloud and its surroundings and tremendously shorten the possible useful life of any artificially-produced cloud. The second limitation results from the impracticability of producing clouds of a thickness comparable to natural clouds that extend vertically thousands of feet. In short, the fundamental limitations imposed by a laboratory are related to the practicable dimensions.

Because the problems of droplet growth and evaporation require a *long free fall trajectory* and because the production of an artificial cloud usually requires the expansion of an air mass from a considerable pressure, it seemed at first that there was little hope of meeting the required conditions without making a prohibitive

FIG. 1. One end of shaft laboratory showing oscillograph and water dropper.

capital investment. It is well established that any mass of saturated air suddenly expanded will cool and produce clouds that are more or less suitable for experimental study. The classical approach employing a moving piston is obviously not practical on a large scale. Therefore, the cyclic compression of the air in a huge chamber with a subsequent sudden expansion has been adopted. In order to induce a sufficient change in temperature to cool the air a few degrees, the fractional pressure change of the air must be considerable. Practically, this implies that the containing vessel can withstand pressures up to at least 5 pounds per square inch throughout its full extent. Thus, the supporting walls of the expansion chamber must be exceptionally strong and well supported. It occurred to the author some years ago that the difficult structural requirements could be met by a vertical mine shaft if the horizontal passages could be suitably isolated from the vertical section. Vertical mine shafts several hundreds of feet deep are relatively common, and the forces arising from the compressional cycle necessary to produce a cloud are passed directly into the adjacent earth-supported walls. Hence, massive supporting structures are not necessary and the cost immediately becomes reasonable. It will be seen that a mine shaft permits the production of a cloud hundreds of feet deep and thus approximates a natural cloud, but the fundamental limitation imposed by the walls still remains. The presence of walls limits the persistence of typical clouds in our shaft to about 15 minutes. Probably this persistence time can be increased to 30 minutes by elaborate thermal treatment of the wall and by the installation of reflecting surfaces having negligible thermal inertia. Such treatment is not contemplated and we are collecting necessary data during the 7- to 10-minute interval when the cloud is changing slowly.

THE SHAFT

The Weather Bureau was indeed fortunate in having the friendly cooperation of the officials of the Miami Copper Company of Miami, Arizona, who made available to the Physical Research Division on a cost basis a fine vertical shaft particularly suitable for our experimental work. The shaft is part of a currently unworked copper mine that once produced more than $300,000,000 worth of copper, silver, and gold. It is now maintained as a valuable source of water and is located on a hill at an altitude of about 3500 feet above sea level. This shaft, 2.2×2.2 meters in cross section and 210 meters deep, is lined with a heavy concrete wall from top to bottom.

Access to the collar of the shaft is through a 150-foot passage to a point about 75 feet below the surface. A cage and hoist were installed in the shaft and these enabled us to carry on with fair efficiency the great labor of preparation necessary to convert the shaft to our purposes. In order to pressurize the shaft, seven heavy bulkheads had to be erected to isolate the shaft from other workings of the mine. Three of these bulkheads are of heavy concrete and steel in which heavy pressurized doors are installed. The other four bulkheads are of heavy wooden and fireproof construction, sealed and fitted to the adjacent concrete abutments in such a way as to reduce air leakage to acceptable values. At the bottom of the shaft, a large working area about 5×10 meters with a concrete floor was included so that the scientists can carry on work inside the expansion chamber during actual compression or expansion. This working area is accessible both from our own shaft and through a heavy airtight door to a passage leading to another cage and hoist 150 feet away.

The passage providing access to the top of the shaft was enlarged at the shaft end and a small laboratory built therein. Necessary shop facilities, benches, and scientific apparatus were installed to expedite the work. This laboratory is about 150 feet from the main surface entrance on the side of the hill and about 75 feet under-

ground. Figure 1 shows one end of the laboratory with the oscillograph and timing clock on the bench.

For certain measurements on the evaporation of freely falling waterdrops, it was necessary to maintain the temperature and relative humidity uniform throughout the shaft. To eliminate wet spots that might vitiate such measurements, all of the concrete joints in the shaft were completely sealed so that surface water would not leak into the shaft or run down the walls. One or two large high pressure leaks resisted damming and had to be sealed and piped off into other parts of the mine. Although sealing the shaft was a big job, it is now *perfectly dry*.

Another great difficulty was to seal the shaft at every point to reduce air leakage resulting from the higher air pressure inside the shaft during the compression period. This laborious task was accomplished by a variety of means including a large amount of patching with mortar and concrete and finally the painting of the entire inside of the shaft with a cement-type waterproof paint. Although the shaft is not yet perfectly airtight, the leakage has been so reduced that more than 30 minutes is now required for the pressure to reduce to $1/\epsilon$ its original value. Such a leakage does not interfere with our operations.

Five 2300-volt armored cables carrying 25- and 60-cycle electric power to a pumping station 200 feet below our lowest level go down the shaft adjacent to our equipment. The electrical interference from these cables has been serious. In particular, the stray currents from the electric power system returning to ground through the outer armor caused great difficulty. Since the water supply for the area depends on the continuous excitation of these circuits, it was necessary to modify our apparatus so it would operate even in the presence of this interference. Telephones are provided to connect the workers on the cage with the laboratory and with the hoist operator.

A photograph taken at the midpoint of the shaft, looking straight up is shown in Fig. 2. One of the inductors for the electronic reporters is shown in the foreground with its screen shield folded back behind the large pipes. In the corner near the reporter are mounted a hydrograph and a thermograph.

In order for the shaft to be of the maximum use, certain air conditioning is necessary and facilities are installed to control, over a limited range, the temperature, the humidity, the pressure, electrical conductivity, and dust content. Refrigeration of the shaft is not considered feasible, and for the time being, at least, our work will be restricted to studies at temperatures above freezing.

In order to modify the temperature and humidity in the shaft, a large steam boiler capable of producing ten million calories per hour has been provided, and a large steam line run from the top of the shaft to the bottom. Vents are provided every 40 feet so that steam can be introduced into the shaft, at will, to humidify

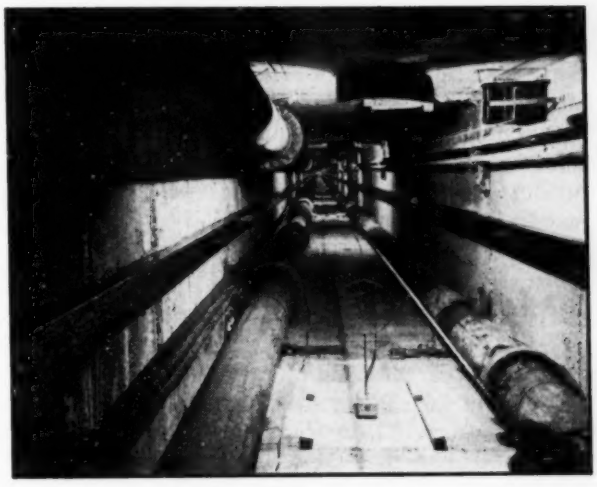


FIG. 2. Looking straight up the shaft from midpoint—reporter in foreground.

the air and moisten the walls of the shaft as required. A water pipe similarly runs from the top to the bottom and is provided with a large number of outlets so that water may be sprayed on the walls or into the air for cooling and humidification. In addition to steam and water, electric power may be dissipated locally, and gas to produce heat or local ionization at a number of positions in the shaft is available in a separate pipe.

The artificial clouds are produced in the shaft by the controlled expansion of the saturated air within. In order to compress the air in the shaft prior to the final expansion, a large compressor has been provided. This compressor of the Roots Blower type is shown in Fig. 3 and is designed to supply 350 cubic feet of air per minute at a maximum of 12 pounds per square inch. It is connected to the shaft through a 4-inch pipe that is also connected to the steam boiler. The air from the compressor is usually at the outdoor temperature and has a low humidity so that the introduction of steam into the compressor line is necessary to preserve the

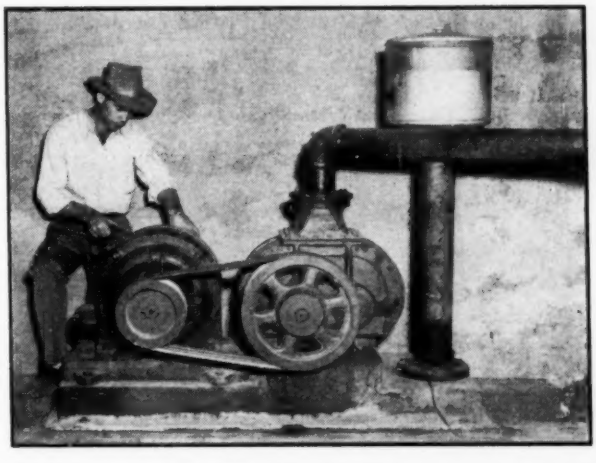


FIG. 3. Compressor for the shaft.

high humidity established in the shaft by other means. Thus, when sufficient steam has been introduced into the shaft to bring it to a suitable temperature and the air is saturated, it is found that the walls are covered with a film of moisture. A number of successive expansions to produce clouds may then be made without the further introduction of steam except in the air compressor line.

ELECTRONIC REPORTERS

To expedite the studies of evaporation and growth of *freely falling* drops, electronic reporters of the type described in earlier communications[1,3] were installed on the walls of the shaft at 10-meter intervals all the way to the bottom. These electronic reporters are made demountable so that the cage of the hoist can go up and down the shaft when they are not in position. Each reporter is capable of producing an amplified pulse that is fed into a concentric electrical transmission line connected to the surface laboratory where a recording oscillograph is located. Means are provided at the top of the shaft to generate in succession waterdrops of any selected size and simultaneously place thereon a free electrical charge sufficiently large to operate the reporters.[3] Whenever a highly charged droplet passes one of the reporters, a pulse is transmitted to the surface that marks the time of droplet passage. Thus, if a small waterdrop is started down the shaft with an electrical charge on it, the time of its arrival at each of the twenty stations is put successively on a moving chart, and this may be compared directly with the time signal traced on the same chart. Using data connecting the velocity of free fall with the mass of the drop, it has been possible to measure the evaporation and growth of drops falling freely down the shaft.

Serious difficulties were encountered in developing the electronic reporters to the point where the electrical

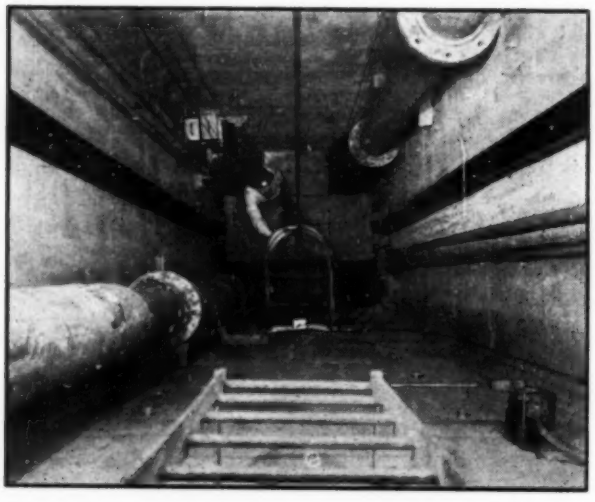


FIG. 4. Looking down the hatch at a worker mounting the reporter at the 10-meter station.

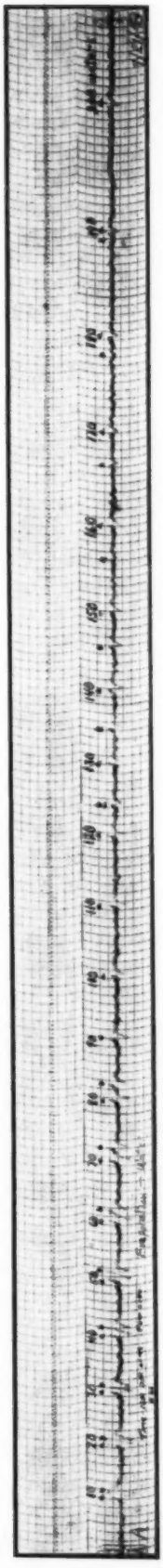


FIG. 5. Actual oscillographic record produced by charged drop falling freely down the shaft.

charges normally placed on the small droplets could be measured against the background of noise in the shaft. The presence of high voltage high power cables introduced noise and interference that could only be eliminated by very careful shielding, not only of the transmission lines and amplifiers, but also the induction electrodes of the reporters. A typical installation at the first 10-meter station below the collar is shown in Fig. 4. A six-foot man standing on top of the cage gives a good idea of the dimensions of the reporters and their electrostatic shields. An actual oscillographic record obtained with this apparatus and exhibiting definite evaporation from such a freely falling drop is shown in Fig. 5.

CLOUD PRODUCTION

Fine, heavy clouds, uniformly filling the shaft from top to bottom, can be produced by saturating the air, as above, and compressing with saturated air up to gauge pressures of about 2.0 pounds per square inch. Since the shaft is at an altitude of 3500 feet, this compression is more or less equivalent to compressing the air to sea level values. When the compressor is shut down, a light cloud is formed immediately because of natural air leakage in the shaft. A 16-inch blow-valve is then opened by a predetermined amount, and the pressure is reduced to $1/\epsilon$ of its initial value in a measured time called the effective expansion time. When this value of pressure is reached, the valve is opened wide, thus reducing the pressure to zero. It has been found that the character of the cloud is determined in an important way by the rate at which the air is expanded. Two fundamentally different types are recognizable. If the cloud is formed in an effective blow-down time of about 10 seconds, an optically very dense cloud is produced in which an observer can see only about 4 meters. If a large number of nuclei are in the shaft, a cloud so produced is composed of cloud elements of a diameter less than 3 microns. The cloud elements are so small, in fact, that it has been impossible for us to collect them on a slide suitable to determine the droplet size distribution. In general, such a cloud does not persist more than 5 to 7 minutes, and the light scattered by the cloud decreases exponentially.

A much different type of cloud is produced when the expansion time is increased or if the condensation nuclei are notably fewer. When such conditions exist, the water condenses apparently on relatively fewer nuclei and has sufficient time to grow into droplets having a reasonably large mean diameter. The light scattered by this type of cloud is persistent for a considerable fraction of the cloud life and then suddenly decreases towards zero. The droplet distribution size for such a cloud is shown in Fig. 6. The meteorologist will recognize this latter type of cloud as one that closely approximates the droplet size distribution in actual clouds. We are, therefore, examining the physical characteristics and behavior of these clouds and have made measurements on the rate of growth of water droplets of various sizes falling through the artificial cloud.

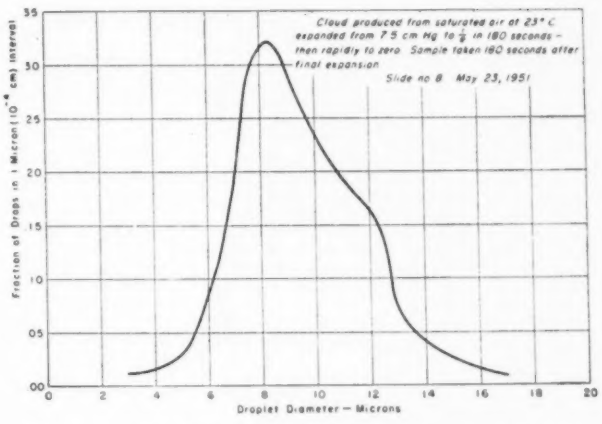


FIG. 6. Droplet size distribution for a "three-minute" cloud.

Our experiments in the new laboratory have already demonstrated that the production, control, and measurement of artificial clouds is entirely practicable. A large number of useful investigations are apparent and it is expected that much new information can be obtained. Data on the evaporation of droplets falling freely down the shaft have been published.[1] We have made substantial progress in the solution of the problems of droplet growth, properties of clouds produced under various conditions, the optical properties of clouds, and a number of other problems that are not yet ready for discussion. We plan to exploit this facility as rapidly as the available support and the small research team working on these matters permit.

Mass-Spectrographic Study of Ionic Reactions in Hydrogen*

RAYMOND L. MURRAY†
Oak Ridge National Laboratory, Oak Ridge, Tennessee
(Received June 21, 1951)

An analysis of the ratios of the di- and tri-atomic ions of hydrogen, deuterium, and H−D mixtures obtained from a hot-cathode arc discharge in a magnetic field was used to examine the nature of ionic collision and displacement processes. Ion mobility, and consequently the rate of ion-molecule collisions, at constant pressure, were controlled by varying the intensity of the magnetic field. Statistical treatment of mass spectrograph data indicates that the formation of H_3^+ is due to the reaction

$$H_2^+ + H_2 \rightarrow H_3^+ + H$$

and that the ion rather than the molecule undergoes dissociation on collision.

INTRODUCTION

STUDIES of the ionization products of hydrogen were summarized by Smyth[1] in 1931. It had been observed that H_2^+ and H_3^+ were more abundant than H^+ and that the H_3^+ increased in relation to the others with increasing gas pressure in the ion source. The origin of tri-atomic ions was not understood, as evidenced by Smyth's remark "the production of H_3^+ remains a great deal of a mystery." He suggested, however, that the main reaction was

$$H_2^+ + H_2 \rightarrow H_3^+ + H$$

instead of the process proposed by Hogness and Lunn[2]

$$H^+ + H_2 \rightarrow H_3^+.$$

Since 1931, the principal interest in tri-atomic ions has been in relation to the analytical determination of isotopic content rather than in the ionic interactions. As late as 1946, Nier[3] reports "the formation is not completely understood, and it has been suggested that the collision of a molecule and a molecular ion in the ionization region produces these anomalous particles."

The discovery by the author that a variation of intensity of the magnetic field in which a hot-cathode hydrogen arc is operated results in a regular change in the H_3^+/H_2^+ abundance ratio suggested a means of examining the processes of ion-molecule collisions. A systematic study was made of the variation of such ratios with magnetic field intensity. The H_3^+/H_2^+ ratio with hydrogen, the D_3^+/D_2^+ ratio with deuterium, and the proportions of the several ion products of masses 1, 2, 3, 4, 5, and 6 with H_2-D_2 gas mixtures were determined. The variations observed could be explained on the basis of three assumptions: (a) that the lighter ions (H_2^+), more strongly constrained by the magnetic field than are the heavier (D_2^+, HD^+), make more collisions in their migration perpendicular to the field; (b) that the reaction $H_2^+ + H_2 \rightarrow H_3^+ + H$ is applicable; (c) that the collision results in the dissociation of the molecular ion rather than the molecule.

It is well known[4] that the arc discharge can be sustained at low pressures in a uniform magnetic field. In particular, a hot-cathode hydrogen arc may be operated at pressures such that the probability of collision of ions with gas molecules in the annulus surrounding the arc column is low. The ion generation rate, however, is large enough to permit the study of primary ions and the collision products with a simple mass spectrograph. In this experiment, the magnetic field serves three purposes: collimation of the ionizing electron stream; control of the number of collisions with the constant pressure gas by virtue of constraint of the moving ions; mass resolution of the ions at 180° from their source. The constraint of ion motion forms the basis of several theories[5–7] of migration perpendicular to a magnetic field. For example, if the radius of circular motion in a free path $r = v/\omega$, where $\omega = eH/mc$, is taken as a criterion of displacement per collision, then the rate of transfer is reduced linearly with increasing field. On the other hand, the amplitude of cycloidal motion in crossed uniform electric and magnetic fields, $a = eE/m\omega^2$, decreases as the square of the field intensity. Ionic diffusion is theoretically described by a modified mobility constant μ_H, related to the ordinary constant μ_0 by

$$\mu_H = \mu_0/(1+\omega^2\tau^2),$$

where τ is the mean time between collisions. For very intense fields, the diffusion rate is again predicted to be inversely proportional to ω^2. On the basis of the mass

* This paper is based on material submitted as a dissertation in partial fulfillment of the Ph.D. degree at the University of Tennessee. Approval and facilities for the investigation were provided by the AEC.

† Now at North Carolina State College, Raleigh, North Carolina.

[1] H. D. Smyth, Revs. Modern Phys. **3**, 147 (1931).

[2] T. R. Hogness and E. G. Lunn, Phys. Rev. **26**, 44 (1925).

[3] Wilson, Nier, and Riemann, *Preparation and Measurements of Isotopic Tracers* (J. W. Edwards, Inc., Ann Arbor, Michigan, 1946), p. 24.

[4] A. Guthrie and R. K. Wakerling, *The Characteristics of Electrical Discharges in Magnetic Fields*, National Nuclear Energy Series I-5 (McGraw-Hill Book Company, Inc., New York, 1949).

[5] J. S. Townsend, *Electrons in Gases* (Hutchinsons Scientific and Technical Publications, 1947), p. 16.

[6] S. Chapman and T. B. Cowling, *The Mathematical Theory of Non-Uniform Gases* (Cambridge University Press, London, 1939), p. 327.

[7] Lewi Tonks and Charles S. Cummings, Phys. Rev. **59**, 514 (1941).

dependence of the angular frequency ω, magnetic constraint would be expected to be most effective for light ions, all other factors being equal.

EXPERIMENTAL EQUIPMENT

The apparatus used in these tests was designed to perform the following functions: to ionize hydrogen gas at low pressure by a stream of high speed electrons; to allow ionic diffusion perpendicular to a uniform (but adjustable) magnetic field, through an essentially gaseous medium; to accelerate and analyze the resultant ions by a mass-spectrograph employing the same magnetic field.

The discharge chamber, Fig. 1, consisted of a 0.5-in. diameter hollow stainless-steel cylinder, 2 in. in length, with a 0.2-in. circular aperture in one end for the admission of electrons from an adjacent incandescent filament. The cylinder was oriented with its axis parallel to the magnetic field. A continuous supply of the gas to be ionized was introduced along one side of the tube, while the ions escaped through a lengthwise slit $\frac{3}{32}$ in. $\times 1\frac{7}{16}$ in. in the other side. The ions were accelerated a distance of $\frac{1}{16}$ in. through another slit, $\frac{1}{16}$ in.$\times 1\frac{7}{16}$ in., into an electric field-free region provided by a "D"-shaped wire mesh cage. A motor-driven ion collector plate, $\frac{1}{16}$ in.$\times 1$ in., was used to scan the 180° plane of this Dempster-type spectrograph, for ion beam diameters from 0.18 in. to 2.3 in. The whole assembly was supported in a vacuum chamber which occupied the space between the pole pieces of an electromagnet. A pressure of the order of 0.1 micron was maintained by mechanical and oil diffusion pumps.

The electrical supplies for the arc consisted of 6-volt rectifiers to provide 350 amperes current for the 0.175-in. diameter tantalum filament, and a 250 volt, 10 ampere capacity rectifier for the filament potential relative to the discharge chamber. The filament heating current was regulated from the arc current, allowing steady arc operation over the periods of time required to take many spectrographic measurements. The filament potential and the arc current were separately adjustable. Ion currents to the scanning electrode were

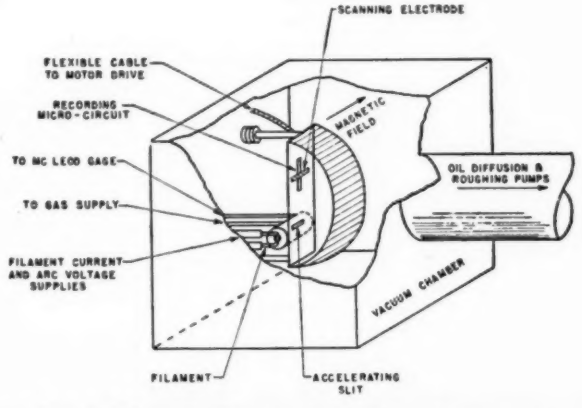


FIG. 1. Schematic diagram of experimental apparatus.

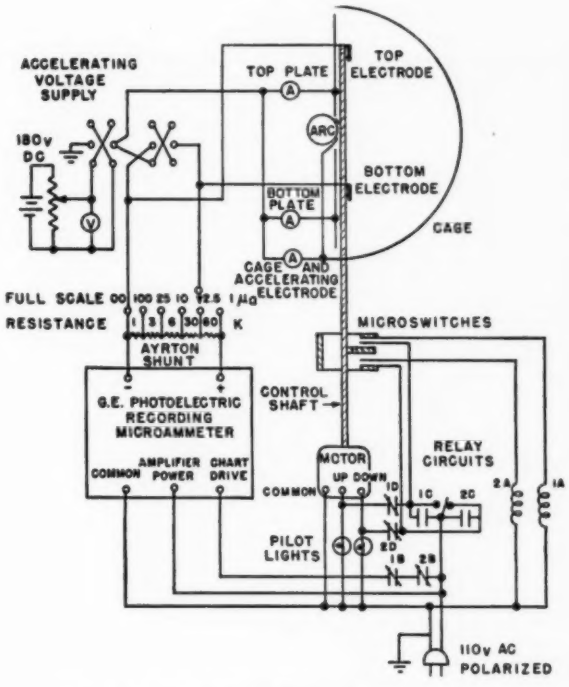


FIG. 2. Control circuit for recording spectrograph.

of the order of 0.1–10 microamperes, measured by a variable-shunt recording microammeter synchronized with motion of the electrode. Each two-inch travel required approximately three minutes. A schematic diagram of the control circuit is shown in Fig. 2. Ion accelerating potentials in the range 0–200 volts, measured with a standard multiscale voltmeter, were provided by dry cells. The magnetic field was measured to within 1 percent and calibrated in terms of the current through the coils of the electromagnet. The mass identification of ions could thus be made with sufficient accuracy for the purpose.

ION REACTIONS IN HYDROGEN AND DEUTERIUM

The first notable observation with hydrogen gas feed, from a commercial cylinder, was that at low electron currents (0–10 ma), low pressure (3 microns), and in the range of magnetic field intensities of 0–1 kilooersteds, the ion abundances of H_2^+ ions and H_3^+ ions were about equal, with H^+ much smaller, Fig. 3. It was also found that the ion proportions changed with field intensity, H_3^+ increasing, apparently at the expense of H_2^+. Although prominent H_3^+ peaks have been observed mass-spectrographically from ordinary arcs,[8] it was at somewhat higher gas pressures than those involved here. The suggested explanation is that the magnetic field constraint reduces the rate of radial transfer, increases the number of collisions, and thus increases the chance of conversion of H_2^+ ions into H_3^+ by a reaction with H_2 molecules. That there is a magnetic constraint on the H_2^+ ions is demonstrated by a se-

[8] H. D. Smyth, Phys. Rev. **25**, 452 (1925).

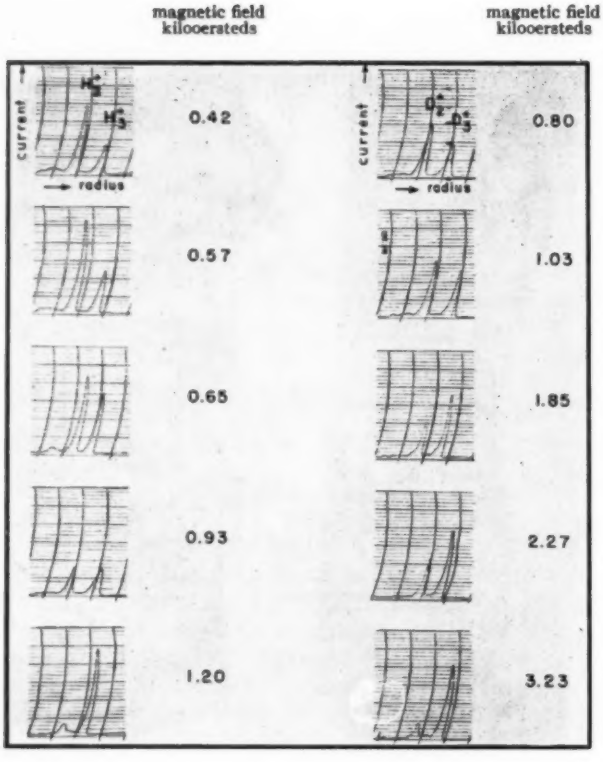


FIG. 3. Spectrograms of ions from H_2 and D_2 discharges for various magnetic field intensities.

quence of spectrograms taken under the constant conditions of arc chamber pressure (2 microns) and electron current (4.5 ma), Fig. 3. The proportions of the two ion types are computed from the heights of the respective current peaks, Fig. 4. The shift with field strength from a preponderance of H_2^+ to that of H_3^+ is evident.

These experiments were repeated with deuterium gas from an electrolytic generator and, as expected, masses 4 and 6, *viz.*, D_2^+ and D_3^+, were in the majority. Figure 3 also shows several spectrograms for deuterium with proportions given in Fig. 4. The difference in the results for deuterium and hydrogen was that the proportion of D_2^+ remains relatively large until much higher fields are reached. One may logically attribute this discrepancy to the relative ease of transfer of masses differing by a factor of two, on the assumption that all other characteristics of the two arcs—electron density, gas pressure, and ionization rate—were the same.

Before attempting to correlate observed data with theory, it is useful to list the best available knowledge of numbers describing the arc system,

dimensions of chamber, $R=0.5$ cm;

mean free path of hydrogen ions in hydrogen,[9] $L=1.43/p$ cm, where p is in microns;

fraction of H_2^+ ions escaping collision in a distance R, neglecting magnetic effects, $f=e^{-R/L}=e^{-0.35p}$;

magnetic angular frequency for H_2^+ ions, $\omega=eH/mc$, 5×10^6 sec^{-1} kilo oersted^{-1};

radii of motion of H_2^+ ions: $H=0.25$ k.o., 1 cm; $H=2$ k.o., 0.15 cm;

velocity of H_2^+ ions of temperature T, in electron volts, $v=10^6(T)^{\frac{1}{2}}$.

The result of the application of a low magnetic field may be considered. At pressures of the order of 1.5 microns with the corresponding mean free path of 0.96 cm, the chance for collision is slight in a distance of 0.5 cm, so the orbit is essentially free. With a magnetic field of the order of 0.5 kilo oersted, and an assumed one volt ion energy, there would be a very mild curvature of the orbits, almost indistinguishable from a straight line insofar as the gas and the dimensions of the vessel are concerned. As the field is increased to 1 or 2 kilo oersteds, however, the radius of curvature decreases to considerably smaller values than the dimensions of the system.

The arc thus appears to be described by the transition region between the cases of "free-fall" and "electric diffusion" in the ordinary low pressure discharge.

ION REACTIONS WITH HYDROGEN-DEUTERIUM MIXTURES

The differences in the effect that magnetic field variation had on the fractions of H_2^+ and D_2^+ in the experiments with pure gases suggested that information on collision phenomena and a test of the assumed mechanism of reaction might be obtained in an investigation of the distributions of ions produced from isotopic mixtures. Two different mixtures were used:

20 percent H_2, 80 percent D_2

59 percent H_2, 41 percent D_2.

The mixed gases were generated by electrolysis of solutions containing the requisite heavy and light water content, account being taken of the hydrogen in the sulfuric acid.

Mass spectrograms for the 20–80 case above, taken

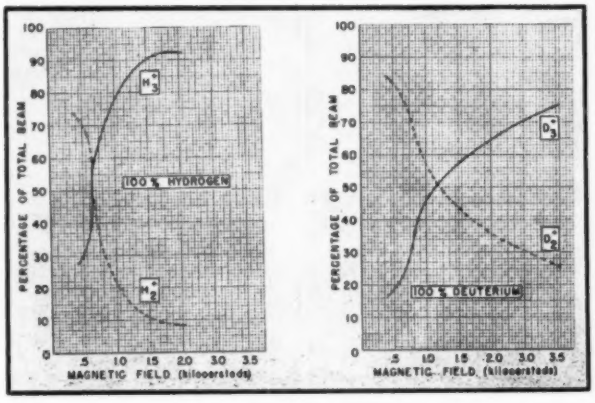


FIG. 4. Proportions of di- and tri-atomic ions from H_2 and D_2 discharges.

[9] A. Ll. Hughes and A. M. Skellett, Phys. Rev. **30**, 11 (1927).

under the following conditions: gas pressure—4 microns; electron current—4 ma; filament potential—100 volts; displayed peaks of masses 1, 2, and 3 as with hydrogen, and masses 2, 4, and 6 as with deuterium, Fig. 5. A prominent peak of mass 5 was new, however. Presumably, it must be only due to HD_2^+ ions formed by some or all of the following collisions: $H_2^+ - D_2$, $D_2^+ - HD$, $HD^+ - D_2$, and $HD^+ - HD$. The HD molecules and ions arise as a result of combination of the electrolysis products. A typical pattern for the 59–41 mixture is also given in Fig. 5. The main point of interest is that the mass 6 peak is now almost missing. The fractions of various types obtained from peak height measurements are plotted in Fig. 6. The ion distribution and its variation with magnetic field had several logical features, such as the decrease in mass 2 and the increase in masses 5 and 6; but also some apparent anomalies, such as the difference in trend of mass 4 for the two cases, and the decrease of mass 3.

ANALYSIS OF FORMATION OF ION TYPES

A systematic analysis of yields on the basis of probability was performed to derive consistency between the data on pure and mixed gases. The factors involved in the problem were classified as follows.

Isotopic Constitution of Gas

The electrolysis process served to prepare three molecular species, H_2, D_2, and HD, in proportions dependent on the concentration of hydrogen and deuterium in the solution. Probability considerations predict that the fractions of these three types obtained in the gas from a mixture of concentrations x and y in H and D, respectively, will be H_2—x^2, HD—$2xy$, D_2—y^2. For example, the gas from 20–80 mixture in solution was calculated to consist of 4 percent H_2, 32 percent HD, and 64 percent D_2, and the gas from the 59–41 mixture 35 percent H_2, 48 percent HD, and 17 percent D_2. The assumption was made that the amounts of hydrogen and deuterium ions liberated at the cathode were equal for equal concentrations in solution, i.e., that there was an inappreciable isotopic separation effect.[10,11]

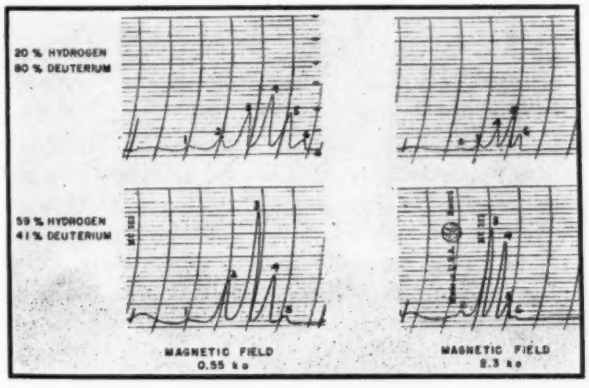


FIG. 5. Typical 180° spectrograph scans from discharge in mixtures of H_2 and D_2.

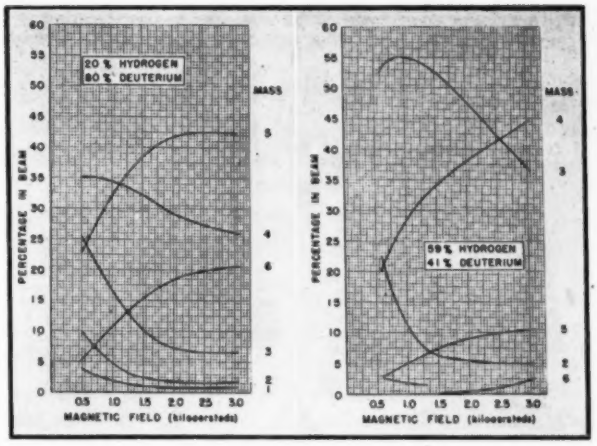


FIG. 6. Experimental variation in proportions of ions in mixtures of H_2 and D_2.

Ionization

The very low H^+ yield would indicate that the predominant primary ionization process was

$$H_2 + e \rightarrow H_2^+ + 2e$$

with corresponding reactions for HD and D_2. Differences in ionization cross section associated with isotopic mass were neglected, being of secondary importance to magnetic field effects. The number of ions of each of the three varieties formed per unit time was taken as proportional to the concentration of each molecular species in the gas.

Collision Rate

Data for differences in the rate of conversion of diatomic to tri-atomic ions have been presented earlier. The probability of a collision of H_2^+ (or D_2^+) ions with another particle to effect an attachment was taken as the fraction of H_3^+ (or D_3^+) ions obtained from the pure gas. Lacking other bases for estimate, the collision probability of HD^+ ions was assumed to be intermediate between those of H_2^+ and D_2^+ ions.

Types of Reactions

The reaction

$$H_2^+ + H_2 \rightarrow H_3^+ + H$$

was taken as the principal event occurring on collisions. In a pure gas reaction the origin of the third atom is indeterminate. Ambiguity also arises in the analogous collision of H_2^+ with D_2, or D_2^+ with H_2. In particular,

[10] H. D. Smyth, *Atomic Energy for Military Purposes* (Princeton University Press, Princeton, 1945), p. 163.

[11] A. Farkas, *Orthohydrogen, Parahydrogen and Heavy Hydrogen* (Cambridge University Press, London, 1935), p. 123.

TABLE I. Schematic analysis of H−D reactions.

An ion of type	Appearing in concentration	Collides (or not) with probability	With a molecule or type	Of concentration	To form the ion type	Of mass
			H_2	X	H_3^+	3
H_2^+	X	B_x	D_2	Y	HD_2^+	5
			HD	Z	H_2D^+	4
		A_x	···	···	H_2^+	2
			H_2	X	H_2D^+	4
D_2^+	Y	B_y	D_2	Y	D_3^+	6
			HD	Z	HD_2^+	5
		A_y	···	···	D_2^+	4
			H_2	X	H_2D^+	4
					H_3^+	3
HD^+	Z	B_z	D_2	Y	HD_2^+	5
					D_3^+	6
			HD	Z	HD_2^+	6
					H_2D^+	5
		A_z	···	···	HD^+	3

two reaction choices are open in the first case:

$$H_2^+ + D_2 \rightarrow H_2D^+ + D$$

$$H_2^+ + D_2 \rightarrow HD_2^+ + H.$$

One may think of these as being, respectively, the dissociation of the D_2 molecule with the attachment of one of its atoms, and the dissociation of the H_2^+ ions and the attachment of its ion. A similar problem arises in the case of the collisions of H_2^+ with HD, D_2^+ with H_2^+ or HD, and HD^+ with H_2 or D_2. To distinguish these cases, recourse may be made to the consideration of relative stability of the molecule and molecular ion. Since the energy of formation of H_2^+ is 60.95 kilocalories per mole and that of H_2 is 102.62 kilocalories per mole,[12] the molecule should be the more stable in a collision with another particle. The reactions in which the molecule remains intact should be favored.

In the event that two products are possible, as in the collisions listed below:

$$HD^+ + D_2 \rightarrow HD_2^+ + D \quad \text{or} \quad D_3^+ + H$$
$$HD^+ + HD \rightarrow H_2D^+ + D \quad \text{or} \quad HD_2^+ + H$$
$$HD^+ + H_2 \rightarrow H_2D^+ + H \quad \text{or} \quad H_3^+ + D$$

equal probabilities were assigned. An attempt to compute the true relative probability would of course be in order.

Four separate probabilities were combined: that of a given ion's presence, its chance of colliding or not (depending on magnetic field), the probability that it reacted with a particular type of neutral molecule, and the possibility of two products. Table I shows a schematic diagram of the algebraic analysis, and Table II gives the total yields of ions of a given mass number, obtained by summing the probabilities of separate modes of formation.

The collision probabilities used were taken from the data for H_2^+ and D_2^+ of Fig. 4. Theoretical trends and relative values, plotted in Fig. 7 for comparison with actual data of Fig. 6, are seen to agree rather well, considering the roughness of the analysis and the uncertainty in both sets of experiments.

TABLE II. Theoretical ion yields.

Mass	Total probability of formation
2	XA_x
3	$XB_xX+ZA_x+\frac{ZB_xX}{2}$
4	$XB_xZ+YA_y+YB_yX+\frac{ZB_xX}{2}+\frac{ZB_xZ}{2}$
5	$XB_xY+YB_yZ+\frac{ZB_xY}{2}+\frac{ZB_yZ}{2}$
6	$YB_yY+\frac{ZB_yY}{2}$

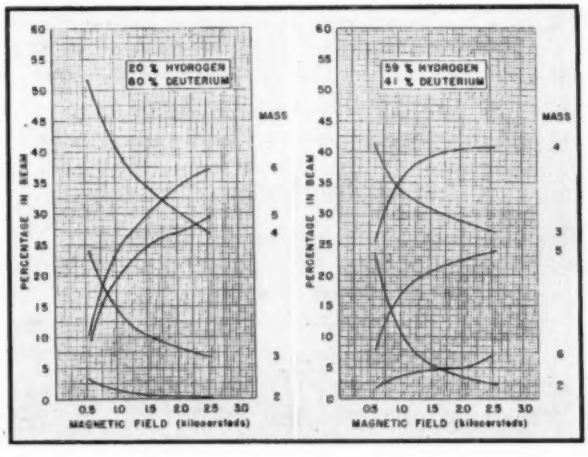


FIG. 7. Theoretical variation in proportions of ions in mixtures of H_2 and D_2.

Conclusions

In the experiments described above, hydrogen, deuterium, or mixtures of the two elements were subjected to an ionizing stream of electrons collimated by an axial magnetic field. The ion products emerging from the gaseous annulus surrounding the stream were analyzed by a simple mass-spectrograph, at various magnetic field intensities. The explanation of the physical reactions and effects appears to coincide with the assumptions made in the analysis of yields:

(a) The proportions of molecular species that enter the arc are governed by the statistical combination rules for free hydrogen and deuterium atoms. (b) The main primary ionization product is H_2^+. (c) There is

[12] Linus Pauling, *The Nature of the Chemical Bond* (Cornell University Press, Ithaca, 1948), pp. 16 and 23.

no appreciable preferential ionization of one of the isotopic molecules H_2, HD, or HD. (d) The cross sections for collision of different isotopic species of ions and molecules are approximately the same. (e) The number of collisions of diffusing ions with gas particles is an increasing function of magnetic field intensity and a decreasing function of ion mass. (f) The principal reaction occurring in hydrogen ion-molecule collisions is $H_2^+ + H_2 \rightarrow H_3^+ + H$. The collision of ions in isotopic mixtures may be represented more generally as $(AB)^+ + (CD) \rightarrow (ACD)^+ + B$ or $(BCD)^+ + A$. (g) In the above reactions the molecular ion is dissociated and the resulting atomic ion combines with the neutral molecule to form the tri-atomic ion. (h) No marked selection of the atomic ion for this attachment is evidenced. Although the comparisons possible with these data do not allow finer distinctions in differences in ion transport or in the mechanism of the collision reaction, the author is convinced that the technique has possibilities of further development into a precise tool for the study of arcs and of atomic collisions. For instance, a more exact measurement of relative probabilities of ion and molecular dissociation on impact could be made.

The author wishes to express his appreciation of the discussions with and the guidance of Dr. E. D. Shipley, Associate Director of the Oak Ridge National Laboratory, to Dr. R. S. Livingston for making available the research facilities and for his encouragement, to Dr. F. T. Howard for editorial suggestions, and to Messrs. J. J. Murphy and J. H. Cupp for technical assistance.

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 1 JANUARY, 1952

An Experimental Investigation of the Dielectric Rod Antenna of Circular Cross Section Excited in Rotationally Symmetrical Modes*

C. M. McKinney
Texas Technological College, Lubbock, Texas
(Received July 9, 1951)

The radiation patterns were measured for three series of dielectric rod antennas of circular cross section excited in the TM_{01} mode at a frequency of 9275 mcps. Only one parameter was changed in each series of rods so that the effect of this parameter on the pattern could be studied. Dielectric antennas excited in the TM_{01} mode have a null in the radiation pattern in the forward direction. It was found that maximum attenuation of secondary lobes was obtained using large diameter rods but sharper and deeper central nulls were obtained with small diameter rods. It was also noted that the deepest central null was obtained using a rod of uniform cross section. For uniformly tapered rods of various lengths, the maximum secondary lobe attenuation as well as the sharpest and deepest central null was obtained using the longest rod. The radiation patterns of the antennas excited in the TE_{01} mode were of the same nature as those obtained using the TM_{01} mode. Measurements were made also of the electric field strength on the surfaces of the radiating rods and these measurements were correlated with the radiation patterns.

I. INTRODUCTION

THE dielectric rod antenna of circular cross section excited in the dominant or HE_{11} mode has been investigated extensively, both theoretically and experimentally, by numerous workers.[1–4] The popularity of the HE_{11} mode is entirely understandable since it is the lowest ordered mode which produces a maximum of radiation in the forward direction. Antennas excited in the TM_{01} and TE_{01} modes have a null in the forward direction. There are, however, several possible uses for antennas with a null in the forward direction. For example, since the null can be made more sharp and distinct than the maximum, the antenna can be used as a direction finding device. If the antenna is directed vertically the radiation pattern is omnidirectional in azimuth.

II. DESCRIPTION OF ANTENNAS; EXPERIMENTAL PROCEDURE

Measurements were made on three series of dielectric rod antennas which have been described in a previous paper.[4] Series A were rods of uniform cross section and length but of various diameters, series B rods were of constant length with different degrees of tapered cross section, while series C rods had a uniform taper but were of various lengths.

The dielectric rods extended into metal guides several wavelengths and were tapered to a point to affect a suitable impedance match. Measurements were made on the rods excited in the TM_{01} and TE_{01} modes. For

* The work described in this paper was done at Defense Research Laboratory, The University of Texas, under the sponsorship of the Bureau of Ordnance, Navy Department, Contract NOrd-9195.

[1] Peter Mallach, "Dielectric directional antennas for dm and cm waves," Air Materiel Command Report, F-TS-2223-RE, February, 1948. Translated by P. L. Harbury of Harvard University, Cambridge, Massachusetts.

[2] G. Mueller and W. Tyrrell, Bell System Tech. J. **26**, 837 (1947).

[3] D. F. Halliday and D. G. Kiely, *Dielectric Rod Aerials*, J. Inst. Elec. Engrs. (London) **94**, 610–618 (1947).

[4] C. W. Horton and C. M. McKinney, J. Appl. Phys. **22**, 1246 (1951).

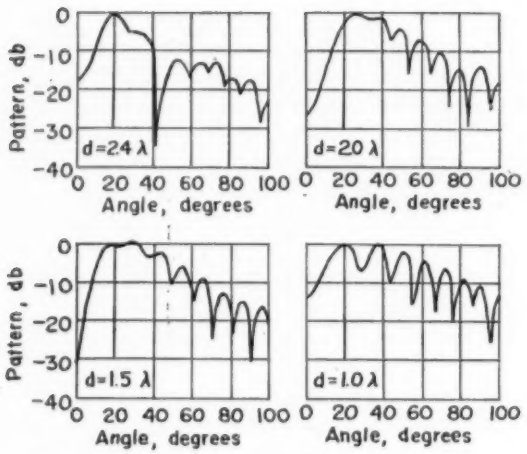


FIG. 1. Radiation patterns for the A-series of rods. Length 6λ; material polystyrene; diameter, d, constant.

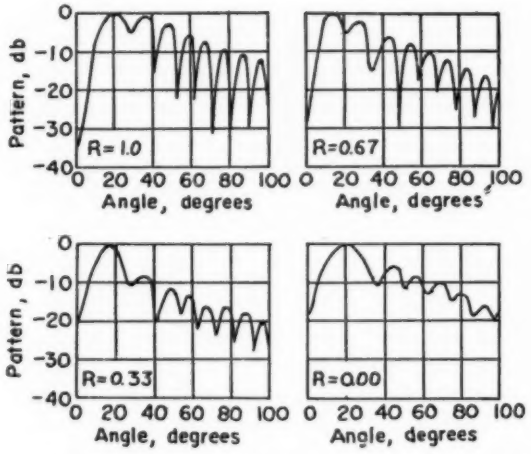


FIG. 3. Radiation patterns for the B-series of rods. Length 6λ; material, Lucite; R=minimum diameter/maximum diameter.

the larger diameter rods the TM_{02} and TE_{02} modes were also present. Characteristics measured were the radiation patterns and the electric fields on the surfaces of the radiating rods. The power source used was a type 723 A/B Klystron operating at a wavelength of 3.20 cm. The receiving system consisted of a metal horn antenna, superhetrodyne receiver, and suitable recorder.

III. EXPERIMENTAL DATA

(a) TM_{01} Mode

The electric and magnetic fields in and about the dielectric rods are similar to the fields on long dielectric wave guides. The magnetic field is circular and transverse. The electric field is predominantly radial and is rotationally symmetrical. At the surface it is almost normal. This type of field naturally produces a null in the forward direction. The radiation pattern is the same for all planes in which the rod lies. The receiving antenna must also be polarized so that it lies in the plane in which the radiation pattern is measured. Figure 1 shows the radiation patterns of the A-series as a function of diameter. As expected, the number of lobes is essentially independent of diameter but the envelope of the radiation patterns is very definitely affected by the diameter.

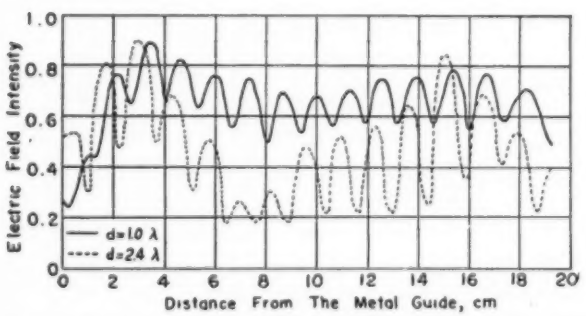


FIG. 2. The normal components of the electric field at the surface of two rods in the A-series. Length 6λ; material, polystyrene; diameter, d constant.

Maximum minor lobe attenuation is obtained with large diameters. From Fig. 1 it appears that the 1.5λ diameter rod has the sharpest and deepest null. Figure 2 shows the normal component of the electric field as a function of axial distance on the surface of two of the A-series antennas. The solid curve is that of a rod 1.0λ in diameter. Only the lowest order mode can propagate, and the field pattern is uniform. The standing wave pattern is due to reflections from the end of the rod. The dashed curve is that of the electric field on the surface of the 2.4λ diameter rod. In this case the diameter of the rod is sufficiently large to support propagation of the TM_{02} as well as the TM_{01} mode.[5] No effort was made to generate any higher order modes but due to the discontinuity at the junction between the metal guide and air, other modes were generated and those were propagated which could be supported by the dielectric guide. Since the wavelengths on the dielectric rod for the two modes were not the same, a "beating" effect occurred and a very nonuniform standing wave pattern resulted. The standing wave ratio was larger for this rod than for the smaller rod because of a greater impedance mismatch at the terminated end.

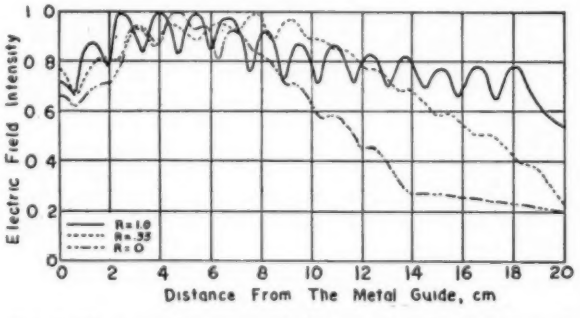


FIG. 4. The normal component of the electric field at the surface of rods in the B-series. Length 6λ; maximum diameter 0.87λ; R=minimum diameter/maximum diameter; material, Lucite.

[5] C. W. Horton and C. M. McKinney, "The dielectric wave guide of circular cross section," to be published in Proc. Inst. Radio Engrs.

Figure 3 shows the radiation patterns of the *B*-series of antennas. Two trends should be noted. A tapered rod appears to produce the maximum attenuation of the secondary or minor lobes although the effect is not noticeably increased for the larger tapers. Of possibly more significance if the antenna is to be used as a null seeking device, is the fact that the nontapered rod has the deepest null and it is almost as sharp as the rod with a ratio of minimum diameter to maximum diameter of 0.66. Figure 4 shows the normal component of the electric field on the surface of three of the *B*-series of antennas measured as a function of axial distance. Again, the field is fairly uniform for the rod of uniform cross section. For the rod with a ratio of minimum to maximum diameter of 0.33 the excitation field decreases rather markedly for the last half of the rod. This effect may be ascribed to the radiation of energy along the preceeding part of the rod, resulting in a gradual decrease of the surface field. The slight increase in intensity along the first half of the rod is associated with a change in phase velocity (which is a function of the diameter) which is accompanied by an increase in the energy exterior to the rod. It should be noted that for the rod which tapers to a point, part of the rod is too small in diameter to support even the TM_{01} mode and hence it contributes nothing to the radiation patterns.

The radiation patterns for the *C*-series of antennas are given in Fig. 5. Here it is noted that the number of lobes are a function of length, as expected, and the minor lobes are attenuated somewhat more by the longer rods. These characteristics are in general the same as for all linear arrays. It should be noted that the minimum central null width and maximum depth was obtained with the longer rod also. Measurements of the surface fields on the rods are omitted for brevity but they were similar to the fields on the *B*-series of antennas. For both the *B*- and *C*-series, the standing wave ratio on the rods was smaller for the tapered rods and was accompanied by a gradual transition of the phase velocity from that on the rod to that of free space.

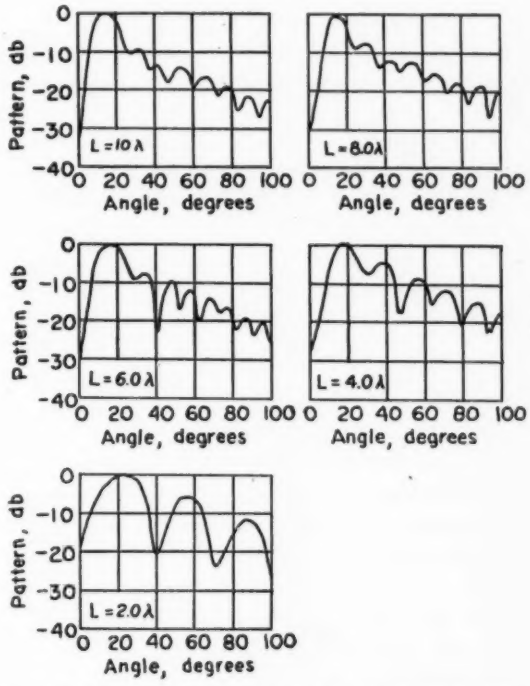


FIG. 5. Radiation patterns for the *C*-series of rods. Maximum diameter 0.87λ; taper 1.04 inches per foot; material, Lucite; length *L*.

(b) TE_{01} Mode

Dielectric rods excited in the TE_{01} mode have the electric field circular and entirely transverse. Again this results in a null in the forward direction. However, in the measurement of radiation patterns, the receiving antenna must be polarized normal to the plane of the radiation pattern. Since the radiation patterns and surface fields were very similar to those obtained when the antennas were excited in the TM_{01} mode, none of the data are included here. The general characteristics and trends were the same for both modes of excitation. The TE_{01} mode appears to have no advantages over the TM_{01} mode and does have the disadvantage of being more difficult to obtain in a pure form.

The Collapse and Rebound of a Gas Bubble*

LEON TRILLING
Hydrodynamics Laboratory, California Institute of Technology, Pasadena, California
(Received June 23, 1951)

The velocity and pressure field in a slightly compressible liquid resulting from the collapse of a spherical bubble is computed as a function of the pressure at the bubble wall, using the acoustic approximation. The results are accurate as long as the liquid velocities are small compared to the sonic velocity in the liquid; they agree with the results previously obtained by C. Herring.

The following bubble model is then investigated. The bubble is supposed filled with inviscid perfect nonconducting gas; special emphasis is given to the gas motion which involves a series of shock waves. A fraction of the energy of compression is thus degraded so that the radius of the bubble after the first rebound is somewhat less than the original radius. The pressure variation at the bubble wall is virtually the same as if the gas were compressed uniformly and isentropically.

INTRODUCTION

THE following paper discusses the motion of a spherical gas bubble immersed in a liquid when the liquid is initially at a higher pressure than the gas, with some account of the shock waves which are produced in the gas as the bubble collapses. In his pioneering paper, Rayleigh[1] idealized the problem by neglecting the compressibility of the liquid and identifying the contents of the bubble with a gas at rest at constant temperature. C. Herring[2] studied the motion of explosion bubbles with arbitrary pressure variation in the bubble, including a first order correction for compressibility of the liquid.

The present paper begins with a general discussion of the motion of the liquid during bubble collapse, which leads to the formula found by Herring.[2] An ideal inviscid gas bubble in a nonconducting liquid is then considered; it is shown that the over-all effect of the gas on the collapse history is virtually the same as if the gas were compressed uniformly and isentropically.

1. The Equations of Motion of the Liquid Surrounding a Collapsing Bubble

A spherical gas bubble of initial radius R_0 is immersed in an infinite volume of liquid having an initial density ρ_{L0} and pressure p_{L0} which is greater than the initial gas pressure $P(0)$ inside the bubble. It is desired to find the pressure and velocity field throughout the liquid as a function of the pressure, $P(t)$, exerted by the gas on the bubble wall during the collapse.

Since the flow is entirely radial and therefore irrotational, the radial velocity u can be written in terms of a velocity potential, φ, as

$$u=-\partial\varphi/\partial r. \tag{1.1}$$

Under the given initial conditions, only diverging spherical waves will occur in the liquid. Hence, as long as liquid velocities are small compared to the velocity of sound c_L, in the liquid, one can expect that the velocity potential will satisfy approximately the acoustic equation for diverging spherical waves:[3]

$$\left(\frac{\partial}{\partial t}+c_L\frac{\partial}{\partial r}\right)(r\varphi)=0. \tag{1.2}$$

In this approximation, any variation of c_L from its initial constant value is also neglected.

The equation for conservation of momentum in radial flow is

$$\partial u/\partial t+u(\partial u/\partial r)+(1/\rho)(\partial p/\partial r)=0, \tag{1.3}$$

which may be integrated between r and ∞ to give

$$-\partial\varphi/\partial t+\tfrac{1}{2}u^2+\int_{p_{L0}}^{p}(dp/\rho)=0, \tag{1.4}$$

since $\varphi=0$, $u=0$, and $p=p_{L0}$ at $r=\infty$. When (1.2) is differentiated with respect to t, and $\partial\varphi/\partial t$ is eliminated by means of (1.4), one obtains the equation of motion in a form convenient for the present purpose:

$$ru(\partial u/\partial t)+(r/\rho)(\partial p/\partial t)+\tfrac{1}{2}c_Lu^2+c_L\int_{p_{L0}}^{p}(dp/\rho)+c_Lru(\partial u/\partial r)+(c_Lr/\rho)(\partial p/\partial r)=0. \tag{1.5}$$

The boundary condition $p=P(t)$ at the bubble wall $r=R(t)$ must now be introduced. If $U(t)$ denotes the velocity of the bubble wall, the rates of change of pressure and velocity at the bubble wall satisfy the relations

$$dP/dt=\partial p/\partial t+U(\partial p/\partial r);\qquad dU/dt=\partial u/\partial t+U(\partial u/\partial r). \tag{1.6a, b}$$

The equation for conservation of mass in radial flow can be written

$$(1/\rho c_L^2)(\partial p/\partial t)+(u/\rho c_L^2)(\partial p/\partial r)+\partial u/\partial r+2u/r=0, \tag{1.7}$$

* This study was supported by the ONR.
[1] Lord Rayleigh, Phil. Mag. 34, 94–98 (1917).
[2] C. Herring, OSRD Report No. 236 (October, 1941).
[3] H. Lamb, *Hydrodynamics* (Dover Publications, New York, 1945), p. 490.

where derivatives of the density have been replaced by corresponding derivatives of the pressure with the help of the relation $c_L^2=dp/d\rho$. By solving the four simultaneous equations (1.3), (1.6a, b), and (1.7), one can express the four partial derivatives of p and u at the bubble wall as

$$(\partial u/\partial t)_R=dU/dt+(U/\rho c_L^2)(dP/dt)+2U^2/R; \quad (1.7a, b)$$

$$(\partial p/\partial t)_R=dP/dt+\rho U(dU/dt);$$

$$(\partial u/\partial r)_R=-(1/\rho c_L^2)(dP/dt)-2U/R; \quad (1.8a, b)$$

$$(\partial p/\partial r)_R=-\rho(dU/dt).$$

Substitution of these expressions into (1.5) yields a differential equation for the bubble-wall velocity:

$$R\frac{dU}{dt}\left(1-\frac{2U}{c_L}\right)+\tfrac{3}{2}U^2\left(1-\frac{4}{3}\frac{U}{c_L}\right)=\frac{R}{\rho U}\frac{dP}{dt}\left(\frac{U}{c_L}-\frac{U^2}{c_L^2}+\frac{U^3}{c_L^3}\right)+\int_{p_{L0}}^{P}\frac{dp}{\rho}. \quad (1.9)$$

It is convenient to change from t to R as the independent variable by means of the differential relation $dR=Udt$. It is also desirable to simplify (1.9) by retaining only terms proportional to unity and to (U/c_L), while omitting terms of higher than first order in (U/c_L). It can be shown that the variation of ρ from its mean value ρ_L is proportional to $(U/c_L)^2$ and hence is negligible in this order of approximation. With these simplifications, (1.9) becomes

$$RU\frac{dU}{dR}\left(1-\frac{2U}{c_L}\right)+\tfrac{3}{2}U^2\left(1-\frac{4}{3}\frac{U}{c_L}\right)=\frac{RU}{\rho_L c_L}\frac{dP}{dR}+\frac{P-p_{L0}}{\rho_L}. \quad (1.10)$$

This equation may be multiplied by $2R^2$ and integrated to give

$$R^3U^2\left(1-\frac{4}{3}\frac{U}{c_L}\right)-R_0^3U_0^2\left(1-\frac{4}{3}\frac{U_0}{c_L}\right)=\frac{2p_{L0}}{3\rho_L}(R_0^3-R^3)+\frac{2}{\rho_L}\int_{R_0}^{R}\left(P(S)+\frac{SU(S)}{c_L}\frac{dP}{dS}\right)S^2dS, \quad (1.11)$$

where S is a dummy variable taking the place of R in the integral. In (1.11) the initial velocity U_0 is that produced by the impulse at $t=0$, when the pressure in the liquid changes discontinuously from p_{L0} to $P(0)$:

$$U_0=\frac{P(0)-p_{L0}}{\rho_L c_L}. \quad (1.12)$$

Equations (1.10) and (1.11) are identical with those obtained by C. Herring,[2] and the analysis given previously results from Kirkwood and Bethe.[4] It is easily verified that when $U/c_L\to 0$ and the gas in the bubble is at constant temperature or entropy the familiar results of Rayleigh[1] and Ackeret[5] are found as special cases. In the more general situation when terms proportional to U/c_L are not neglected, (1.11) may be solved for $U(R)$ by successive approximation when the gas-pressure variation $P(R)$ is known; an approximate velocity function $U(S)$ is introduced into the right-hand side of (1.11), and the equation is solved for $U(R)$ to obtain a better approximation; the process is repeated until the desired accuracy is attained. In the most general situation, however, $P(R)$ will not be known beforehand, since the gas pressure may depend upon the unknown bubble-wall velocity. In this situation, (1.11) must be solved simultaneously with the equations for the gas pressure to be derived subsequently.

After $U(R)$ is computed by one of these methods, the velocity potential throughout the liquid can be found from the approximate expression,

$$\varphi(r,t)=\frac{1}{r}\int_{R_0}^{R(t-r/c_L)}\left(\frac{P(S)-p_{L0}}{\rho_L U(S)}+\frac{U(S)}{2}\right)S\,dS=\frac{F(t-r/c_L)}{r}. \quad (1.13)$$

This expression for φ satisfies the wave equation, (1.2), and it also satisfies (1.4) approximately at the bubble wall. The velocity and pressure throughout the liquid can then be found from (1.13) and (1.4):

$$u(r,t)=-\frac{\partial\varphi}{\partial r}=\frac{F}{r^2}+\frac{F'}{c_L r}; \quad (1.14)$$

$$p-p_{L0}=\frac{\rho_L F'}{r}-\frac{\rho_L(c_L F+rF')^2}{2c_L^2r^4}. \quad (1.15)$$

2. The Adiabatic Collapse and Rebound of a Gas Bubble

Let the bubble considered in the previous section be filled with a perfect, nonconducting, nonviscous gas initially at rest at a pressure $p_{g0}=P(0)$ and density ρ_{g0}. An approximate solution for the subsequent gas motion may be obtained with the help of an r, t diagram, as shown in Fig. 1. At the instant $t=0$, the pressure discontinuity $p_{L0}-p_{g0}$ is replaced by a velocity impulse into the gas. Since the separation surface moves into the gas, a shock wave $s_g^{(1)}$ travels ahead of it toward the center of the bubble, while an expansion wave $s_L^{(1)}$ radiates into the liquid. Ahead of these two waves, the regions I_g and I_L are undisturbed. At the instant t_1, the shock wave $s_g^{(1)}$ is reflected at the center of the gas bubble; at t_2, the reflected shock $s_g^{(2)}$ reaches the separa-

[4] J. G. Kirkwood and H. Bethe, OSRD Report No. 588 (1942).

[5] J. Ackeret, "Experimentelle und Theoretische Untersuchungen uber Hohlraumbildung im Wasser," Eidgenöss. Materialprüffungsanstalt E.T.H., Zurich (1930).

tion streamline S, where it is reflected into the gas as a shock $s_g^{(3)}$ and refracted into the liquid as a shock $s_L^{(2)}$. The flow of the liquid in region II_L bounded by the expansion wave $s_L^{(1)}$, the separation streamline S, and the refracted shock $s_L^{(2)}$ is determined with the help of (1.11), (1.13), and (1.14). The flow of the gas in region II_g bounded by the shocks $s_g^{(1)}$ and $s_g^{(2)}$, and by the streamline S, and in region III_g bounded by the shocks $s_g^{(2)}$ and $s_g^{(3)}$, is obtained by solving the nonisentropic gas equations approximately. When all conditions ahead of the shocks $s_g^{(3)}$ and $s_L^{(2)}$ are known, conditions at t_2 are determined, and the point t_2 on S serves as a new initial point.

The flow of the gas is thus divided into triangular regions $I_g, II_g, \cdots, N_g$, where conditions are determined by approximate calculation. All jump conditions across the shocks $s_g^{(n)}$ near the streamline S can be found if the density of the liquid is much greater than that of the gas, and the shocks are weak.

At any point t_{2n} where a shock $s_g^{(2n)}$ meets the separation streamline S, a wave is reflected into the gas, and another is refracted into the liquid. The velocity change across the waves $s_g^{(2n)}$ and $s_g^{(2n+1)}$ is

$$U_{2n+2}-U_{2n}=\frac{1}{\rho_{g2n}c_{g2n}}\left(\frac{2\gamma P_{2n}}{(\gamma+1)P_{2n+1}+(\gamma-1)P_{2n}}\right)^{\frac{1}{2}} \times\left[(P_{2n+1}-P_{2n})-(P_{2n+2}-P_{2n+1}) \times\left(\frac{(\gamma+1)P_{2n}+(\gamma-1)P_{2n+1}}{(\gamma+1)P_{2n+2}+(\gamma-1)P_{2n+1}}\right)^{\frac{1}{2}}\right], \quad (2.1)$$

where c_g is the local sonic velocity, and γ is the ratio of specific heats of the gas. The velocity change across the refracted shock $s_L^{(2n)}$ is approximately

$$U_{2n+2}-U_{2n}=\frac{P_{2n+2}-P_{2n}}{\rho_L c_L}. \quad (2.2)$$

Since velocity and pressure are continuous both immediately before and immediately after t_{2n}, and since the density of the liquid is much higher than that of the gas, the pressure relation across the shocks becomes

$$P_{2n+2}=P_{2n+1}\left(\frac{(3\gamma-1)P_{2n+1}-(\gamma-1)P_{2n}}{(\gamma-1)P_{2n+1}+(\gamma+1)P_{2n}}\right). \quad (2.3)$$

Equation (2.3) is identical with the pressure relation for the reflection of a shock at a solid wall.

Since the shocks involved are weak ($\Delta P/P\sim.10$), one may assume, following Guderley,[6] that after reflection the fluid is at rest, at constant density. It follows that

$$\rho_{2n+1}=\rho_0\left(\frac{R_0}{R_{2n}}\right)^3, \quad (2.4)$$

and the pressure jump is

$$\frac{P_{2n+1}}{P_{2n}}=1+\frac{\gamma(\gamma+1)}{4}\left(\frac{U_{2n}}{c_{g2n}}\right)^2 -\frac{\gamma U_{2n}}{c_{g2n}}\left(1+\frac{\gamma+1}{8}\frac{U_{2n}^2}{c_{g2n}^2}\right)^{\frac{1}{2}}. \quad (2.5)$$

From the shock conditions and (2.4) one finds that the pressure rise from t_{2n} to t_{2n+2} is

$$\left(\frac{\bar{P}}{P}\right)_{2n}=\left(\frac{R_{2n}}{R_{2n+2}}\right)^{3\gamma}\left[1+\frac{U_{2n+2}}{c_{g2n+2}}+\frac{U_{2n}}{c_{g2n}}+\cdots\right]^{\gamma}. \quad (2.6)$$

The time required for the shock $s_g^{(2n+1)}$ to reach center, be reflected, and reach S again is

$$t_{2n+2}-t_{2n}=\frac{R_{2n+2}}{c_{g2n+2}}\left(1+\frac{\gamma+1}{4}\frac{U_{2n+2}}{c_{g2n+2}}\cdots\right) +\frac{R_{2n}}{c_{g2n}}\left(1+\frac{\gamma+1}{4}\frac{U_{2n}}{c_{g2n}}\right)+\cdots. \quad (2.7a)$$

Meanwhile, the separation streamline S has moved inward by a distance

$R_{2n}-R_{2n+2}$:

$$t_{2n+2}-t_{2n}=\int_{R_{2n}}^{R_{2n+2}}\frac{dR}{U}\approx 2\frac{R_{2n+2}-R_{2n}}{U_{2n+2}+U_{2n}}. \quad (2.7b)$$

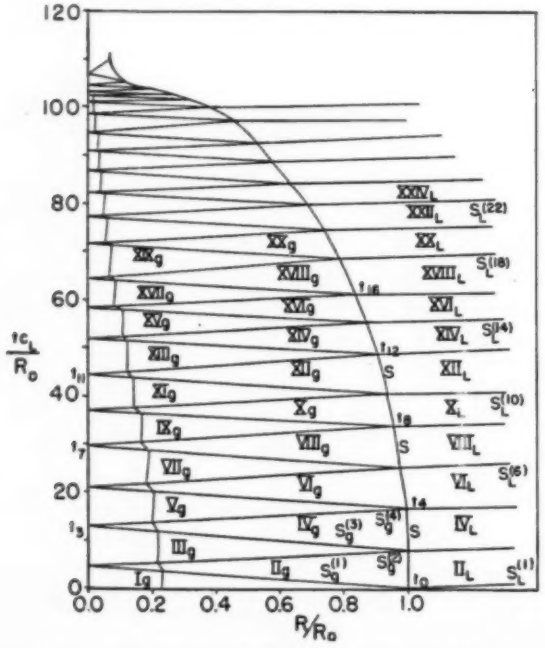


FIG. 1. R, t diagram of the collapse of a gas bubble.

[6] G. Guderley, Luftfahrt-forsch. 19, No. 9, 302–312 (1942).

Therefore, since the shocks in the gas are weak,

$$\frac{R_{2n+2}}{R_{2n}}=1+\frac{U_{2n+2}}{c_{g2n+2}}+\frac{U_{2n}}{c_{g2n}}, \tag{2.8}$$

and the pressure increase along S from t_{2n} to t_{2n+2} is approximately

$$\left(\frac{\bar{P}}{P}\right)_{2n}=\left(\frac{R_{2n}}{R_{2n+2}}\right)^{2\gamma}. \tag{2.9}$$

The shock relations (2.3) and (2.5) and Eqs. (2.8) and (2.9) show the effect of the gas flow on the separation streamline S. Their derivation is based on the assumption valid in acoustic theory and made plausible by Guderley's work, that the velocity behind the reflected spherical shock is small compared with the velocity ahead of it, and that the gas behind the reflected shock is at sensibly constant density.

If the preceding results are substituted into the velocity equation (1.11) and the required integrations are carried out, the following result is obtained:

$$U^2\left(1-\frac{4U}{3c_L}\right)=U_{2n}^2\left(1-\frac{4}{3}\frac{U_{2n}}{c_L}\right)\left(\frac{R_{2n}}{R}\right)^3$$
$$+\frac{2}{3}\frac{P_{L0}}{\rho_L}\left[\left(\frac{R_{2n}}{R}\right)^3-1\right]-\frac{2P_{2n}}{(3-2\gamma)\rho_L}$$
$$\times\left(1-2\gamma\frac{U_{2n}}{c_L}\right)\left[\left(\frac{R_{2n}}{R}\right)^3-\left(\frac{R_{2n}}{R}\right)^{2\gamma}\right]. \tag{2.10}$$

The preceding theory has been applied to compute the collapse of a gas bubble in water ($c_L=1.45\times10^5$ cm/sec, $\rho_L=1$ gm/cm^3); the bubble having an initial radius of 1 cm, initial gas pressure of 0.020 atmos, gas density of 2.6×10^{-5} gm/cm^3, and $\gamma=1.4$; under an external pressure of 1.50 atmos.

The history of the collapse is shown in Fig. 1, which is an r, t diagram showing the position of the separation streamline as a function of time, and the various shock waves. Figure 2 is a plot of the pressure as a function of bubble radius. A minimum radius of 0.064 cm is reached after 0.78 millisec. The pressure maximum is approximately 2200 atmospheres.

After collapsing, the bubble rebounds. The energy available for that expansion is the reversible part of the work done on the bubble:

$$W_{\text{rev}}=(4/3)\pi R_0^3\frac{\rho_{g0}c_{g0}^2}{\gamma(\gamma-1)}\left[\left(\frac{R_0}{R_{\min}}\right)^{3(\gamma-1)}-1\right]. \tag{2.11a}$$

That energy is used to do work on the liquid; the expansion is nearly isentropic. It lasts until all the energy (2.11a) is used up. This condition defines the upper limit

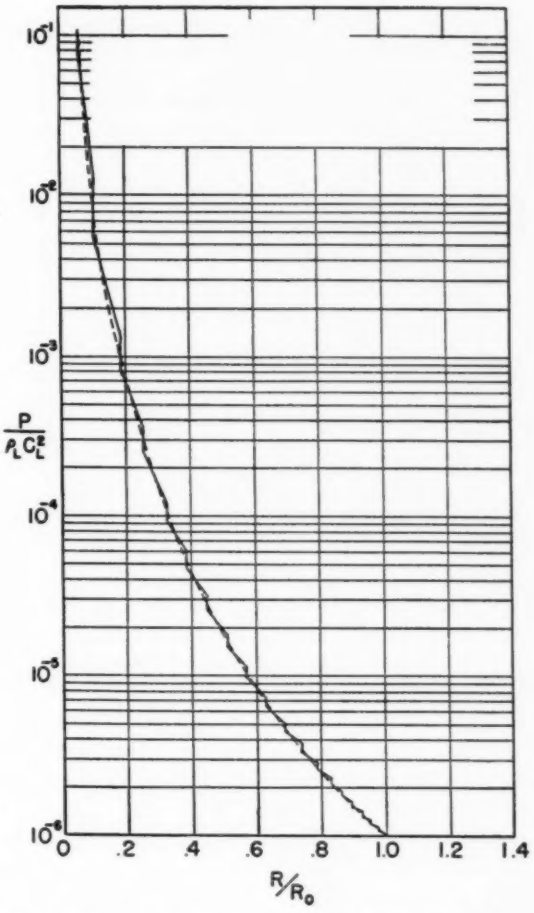


FIG. 2. Pressure distribution on the wall of a collapsing gas bubble compared to isentropic collapse.

of the work integral:

$$W_{\text{rev}}=W_{\text{exp}}=(4/3)\pi R_0^3 P_{\max}$$
$$\times\int_{R_{\min}/R_0}^{R_1/R_0}\left(\frac{P}{P_{\max}}\right)S^2dS. \tag{2.11b}$$

When numerical values for the particular bubble described above are substituted into (2.11), it is found that the rebound radius R_1 satisfies the equation:

$$\frac{(R_0/R_{\min})^{3(\gamma-1)}-1}{(R_0/R_{\min})^{3(\gamma-1)}-(R_0/R_1)^{3(\gamma-1)}}=1.013, \tag{2.12a}$$

whose solution is

$$\frac{R_1}{R_0}=0.85. \tag{2.12b}$$

ACKNOWLEDGMENTS

This study was supported by the Office of Naval Research. The writer wishes to thank Professor M. S. Plesset for suggesting the investigation of the problem, and Dr. F. R. Gilmore for several helpful discussions.

On the Empirical Law of Adhesive Wear

J. T. BURWELL* AND C. D. STRANG†
Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts
(Received July 2, 1951)

Careful measurements have been made of the wear of rubbing steel surfaces as a function of load, distance of travel, and hardness under controlled conditions which eliminated the effect of all factors except adhesion. At normal pressures averaged over the apparent contact area which are less than one third the hardness of the softer material the amount of wear was found to depend linearly on the distance of travel and on the load but was independent of the apparent area of contact. At average pressures greater than one third of the hardness, the depth of wear varied linearly with the distance of travel, was independent of the apparent area of contact for a given average normal stress, but increased many fold for small increases in stress. These findings rationalize the familiar observations on the running-in of machine surfaces. The results are discussed in the light of the current adhesion theory of dry friction.

INTRODUCTION

THE state of our knowledge about wear is in sharp contrast to that about friction. The empirical laws of friction, known as Coulomb's laws, giving the dependence of the friction force on the operating quantities of load, speed and area of contact, have been known for over two hundred years. This is not true of wear. Although the technical literature is full of reported data on wear, these are in almost all cases the results on particular industrial materials having complicated and uncertain compositions run in tests designed to simulate specific operations in service. As a result, it has not been possible to deduce any general laws of wear. In fact, there are only two qualitative conclusions concerning wear to which those familiar with the field would at all agree:

(a) Wear increases in a general way with distance of travel or time of running, but generally not in a linear fashion. The most usual relation is qualitatively as sketched in Fig. 1 but there can be great variations from this form of curve in individual cases.

(b) Wear generally decreases with increasing hardness of the rubbing surfaces but here again many exceptions to this rule can be cited.

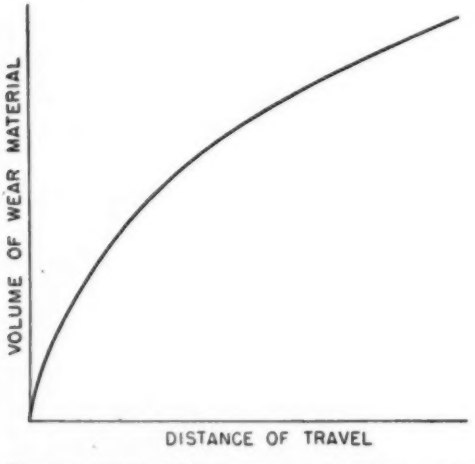


FIG. 1. Typical relation between wear and distance of travel for actual machine parts, especially when freshly assembled.

Dependence on load, area of contact, and other factors are even less generally agreed on.

This lack of any empirical laws of wear makes it impossible to define any "wear coefficient" in the sense that Coulomb's first law enables us to define a coefficient of friction, or Newton's law defines a coefficient of viscosity, or Hooke's law a Young's modulus. As a result, it is obviously not yet possible to apply dimensional analysis to wear problems. This means that at the present time one cannot rationally estimate the wear occurring in an actual machine nor can one conclude anything about it from scaled experiments.

In spite of our lack of knowledge of any relation between the amount of wear and the simple operating variables which will be necessary in order to deduce precisely what the physical mechanism is, still we do know that there are a number of factors which may play a role in the wear process although their relative importance is still in doubt. These factors are:

(1) Adhesion or galling. This is the factor that has been described by Holm,[1] Bowden and Tabor,[2] and Merchant[3] as being largely responsible for friction.

(2) Corrosion.

(3) The presence of loose abrasive material.

(4) The cutting or ploughing of a soft surface by a hard, rough surface much as a file removes material from the work-piece. This has been pointed out by Bowden and Tabor[4] also, to contribute to the friction in certain cases.

(5) A number of less usually occurring factors such as erosion, surface fatigue, etc.

The above factors are arranged roughly in order of increasing avoidability. That is, either with care in practice or under controlled laboratory conditions it is possible to eliminate or at least greatly minimize all of

* Now at Horizons Incorporated, Cleveland, Ohio.
† Now at the Kiekhaefer Corporation, Cedarburg, Wisconsin.

[1] R. Holm, Wiss. Veröffentl. Siemens-Werken **17**, 43 (1938); **20**, 68 (1941).
[2] F. P. Bowden and D. Tabor, Proc. Roy. Soc. **A169**, 391 (1939).
[3] M. E. Merchant, J. Appl. Phys. **11**, 230 (1940).
[4] F. P. Bowden and D. Tabor, Nature **150**, 197 (1942); J. Appl. Phys. **14**, 80 (1943).

them except the first one, namely, adhesion or galling. It has been found that whenever two surfaces are in solid rubbing contact some adhesion and resulting transfer of material always takes place, even in the presence of the best boundary lubricants. This has been demonstrated by the technique of radioactive tracers[5] where one of the two rubbing surfaces is made radioactive and the other surface is examined for evidences of radioactivity indicating that material from the first surface has actually adhered to it. This is illustrated in Fig. 2 both in the presence of an indifferent lubricant, white mineral oil, and a good boundary lubricant, oleic acid. In the latter case the amount of material transferred (tracks located at *B* and *d*) although small is definite and clearly visible on the original negative.

In the light of the above it seemed best to begin by investigating the laws of galling wear as being the irreducible minimum or base wear. If laws for this case can be found, then it may be possible later to deduce the effect of the other factors such as corrosion, etc., by the increment to the basic wear rate which they produce when present. This is the purpose of the present paper.

PREVIOUS THEORY

As a matter of fact, a semi-empirical theory of galling wear has been proposed by Holm[6] although it has not been tested by him or anyone else to any extent. This theory makes use of one of the concepts which has been described as playing a role in the mechanism of friction. This is the idea of the true area of contact between rubbing surfaces, given by the equation

$$A=\frac{W}{p_m}. \tag{1}$$

(W is the normal load and p_m is a flow pressure of the softer of the two surfaces), the physical mechanism being the plastic deformation of the local asperities on the softer surface.

Holm next assumed, that during sliding for every encounter of a surface atom in this true contact area with an atom in the other surface, there was a statistically constant probability of one of the two atoms being pulled out of its parent surface.‡ After counting up all the encounters during the sliding and making use of average interatomic spacings, it turned out rather simply that the volume V of material removed on

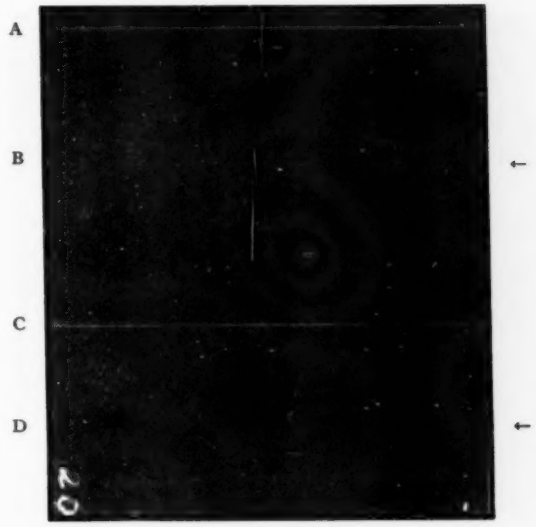


FIG. 2. Autoradiograph of four lubricated friction tracks. Chromium sliding on steel. Lubricant for tracks A and C—white mineral oil; for tracks B and D (very faint)—oleic acid.

sliding a distance L was given by

$$V=k\cdot A\cdot L, \tag{2}$$

where k is the probability of removing an atom.

Eliminating A between Eqs. (1) and (2) gives

$$V=k\cdot\frac{W\cdot L}{p_m}, \tag{3a}$$

as the desired relation. The probability k should be characteristic of the composition of the surfaces, condition of lubrication, etc. For instance if $k\approx 10^{-8}$, it means that on the average each atom in the true contact area has dislodged one atom from the other surface in sliding a distance of 1 cm.

This equation can be written in another form by dividing both sides by the apparent or nominal area of contact Ao to give the average depth of material removed,

$$h=k\cdot\frac{P\cdot L}{p_m}, \tag{3b}$$

where P is the average normal stress over the nominal contact area. Equations (3a) and (3b) may be used interchangeably.

Now it should be noted that in deriving the above, Holm assumed the material to be removed as individual atoms. Actually, we know that a good deal of it is removed as relatively large aggregates of atoms. This is illustrated in Fig. 3 which is an electron microscope picture of the wear particles removed from two hardened steel surfaces rubbing together under a lubricant. The size of the individual specks or grains in the picture, of which the larger black regions seem to be agglomerates, is 50 to 100A which, while quite small, still consist of a million or so atoms. With softer steel the particles are

[5] Sakmann, Burwell, and Irvine, J. Appl. Phys. **15**, 459 (1944).
[6] R. Holm, *Electric Contacts* (H. Geber, Stockholm, 1946), p. 214; *Mechanical Wear* (Am. Soc. Metals, 1950), p. 317.
‡ Strictly speaking, Holm divided the volume of wear material by the true contact area to obtain the average number k of abraded atoms per atomic encounter. He later imagines these abraded atoms as spread in a layer of uniform thickness over the contact area and since he calculates from his data thicknesses less than one atomic layer in some instances, then in those cases k can only be interpreted as a probability of atomic removal. Its constancy, independent of W and L, is implicit in Eq. (3a).

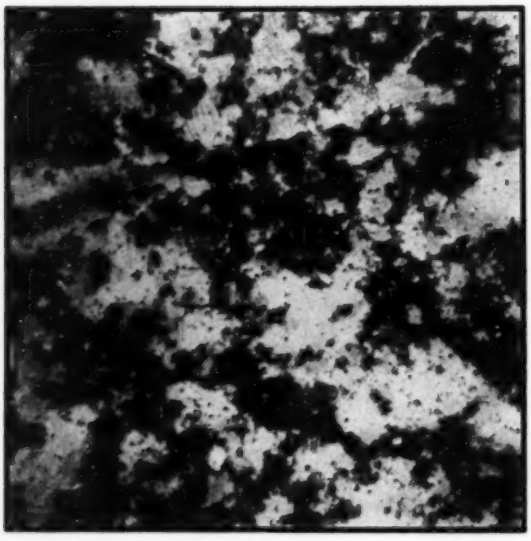


FIG. 3. Electron micrograph of wear particles from hardened steel surfaces (25,000×).

found to be considerably larger. Hence it would seem that Holm's assumption is not generally valid. However, in the absence of any other picture it seemed worthwhile to begin the investigation by testing under controlled conditions to what extent Eqs. (3) holds.

For the sake of completeness we should add the relation between friction force F and true contact area which has been developed by Holm,[1] Merchant,[3] and Bowden and Tabor,[4] namely that

$$F=s\cdot A, \tag{4}$$

where s is the average shear strength of the local adhesions between the surfaces. Dividing Eq. (4) by Eq. (1) gives the familiar expression for the friction coefficient

$$f=\frac{F}{W}=\frac{s}{p_m}. \tag{5}$$

The terms s and p_m are both plastic properties of the material concerned and may be expected to have roughly the same ratio for both hard and soft materials.

EXPERIMENTAL

The apparatus used for the wear experiments was of very simple design. In essence it consisted simply of the end of a relatively soft pin rubbing against the flat face of a smooth, hardened steel disk in a circular track. The disk had a hardness of 550 Brinell and a roughness of 1 microinch rms. The pin was made of a high carbon (SAE 1095) steel which was quite uniform in structure and was of several shapes, cylindrical, conical, and spherical. It was in two different hardnesses and structures, normalized with 223 Brinell and sorbitic with 430 Brinell. The wear depth was determined by measuring optically the diameter of the apparent wear area on the cone or sphere. The wear volume was then calculated from the geometry. In the case of the cylinder where the wear area did not change with depth, it was necessary to measure the length of a shallow indentation previously made in the end of the cylinder after the method of McKee.[7] The area of the indentation was only a very small fraction of the total wear area.

It was considered important to be able to measure the wear depth during a run with a minimum of disturbance of the two rubbing parts because, as will be seen from the results obtained, freshly assembled parts always exhibited a different wear rate from that obtained later in the run. This is the familiar phenomenon of "running-in." In order to do this, the pin was mounted in a very rigid support which had only one degree of freedom so that it could be rotated about a pivot from its running position to the stage of a microscope for reading the dimensions of the wear area or indentation as shown in Fig. 4. The pivot consisted of precision-grade ball bearings, preloaded axially. In this way the pin could be returned quite accurately to its former position on the disk after a reading had been taken. The depth of wear on the cylinder could be measured to 5×10^{-6} cm and that of a 60° cone to 8×10^{-5} cm. In a few runs where it was desired to measure the friction and wear simultaneously, a special support for the pin was used which deflected sufficiently under the action of the friction force to produce signals from four series-connected electric resistance strain gauges mounted on a dynamometer ring against which it bore. The system permitted no tilting or other deflection of the rider so that the rider's wearing surface remained always parallel to itself. A photographic view is shown in Fig. 5 of the apparatus with the rider in the reading position and including the friction-measuring system.

It was also important to eliminate all the factors such as corrosion, loose abrasive, etc., so as to leave only adhesion to produce wear. After some investigation, the

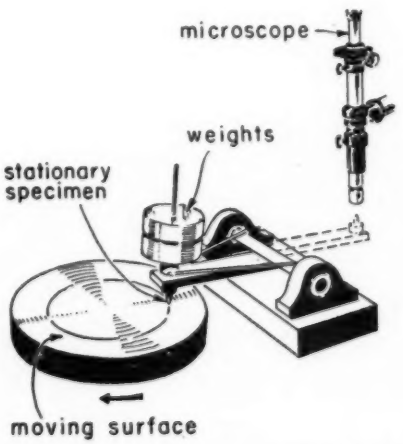


FIG. 4. Sketch of wear apparatus. Full view shows rider in running position, dotted view shows it in position for visual measurement.

[7] S. A. McKee, J. Research Natl. Bur. Standards **39**, 155 (1947).

simplest way of doing this was to flood the disk surface, which was horizontal, with a thin layer of a chemically inert liquid. Surface tension kept the liquid on the rotating disk. In this way the rubbing contact was kept completely immersed at all times and air and grit were excluded. It was found to be important to exclude the air because when the surfaces were run dry a great deal of brown iron oxide was formed indicating corrosive action of the air. The disk surface was cleaned each day with solvent and fresh liquid applied to remove any accumulated grit or wear particles. Purified hexadecane was used as the liquid so as to exclude the humidity from the air, not be too volatile and yet provide a minimum of boundary lubrication. Even so, it was found that on days of high humidity (over 20 percent) the wear rate was markedly higher and brown oxide formed even under the oil. Hexadecane is known to dissolve a small but finite amount of water. Such runs had to be discarded.

The question might well be raised as to why for a "standard state" of wear the experiments were not carried out dry in an inert atmosphere thus eliminating all effect of lubricant. At first glance this sounds reasonable but in practice the question of selecting just what these conditions should be resolves itself into deciding just how completely all nonmetal materials such as residual oxide films, etc., which could act as lubricants, should be removed. If they are not completely removed before the run starts, then the running conditions are transient as such films are gradually worn through. After they are worn through or if they are removed beforehand, violent welding, transfer, and seizure takes place under practically zero load. The situation is quite similar to that of measuring the friction between metal surfaces in vacuo. Here also, although the ambient condition would seem to be the ideal reference one, yet in actual fact friction coefficients of any value either positive or negative, may be obtained at will because of the violent welding of the two surfaces. Such experimental conditions have little significance beyond demonstrating that perfectly clean metal surfaces will weld completely when brought into contact. For this reason running in an inert gas atmosphere can be expected to throw little light on the phenomenon of wear as ordinarily encountered in practice and it was, therefore, decided to use an inert fluid lubricant.

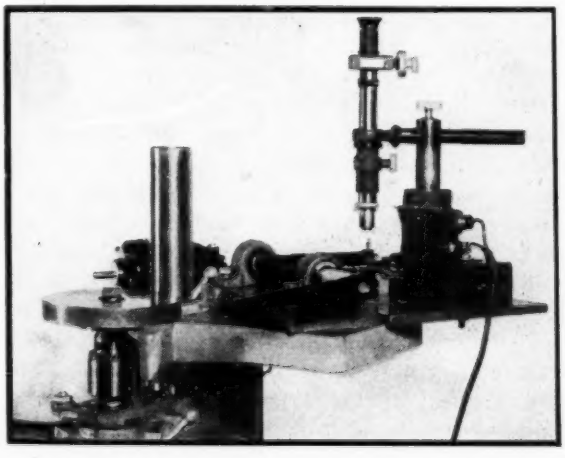


FIG. 5. View of wear apparatus with rider in retracted position for reading wear.

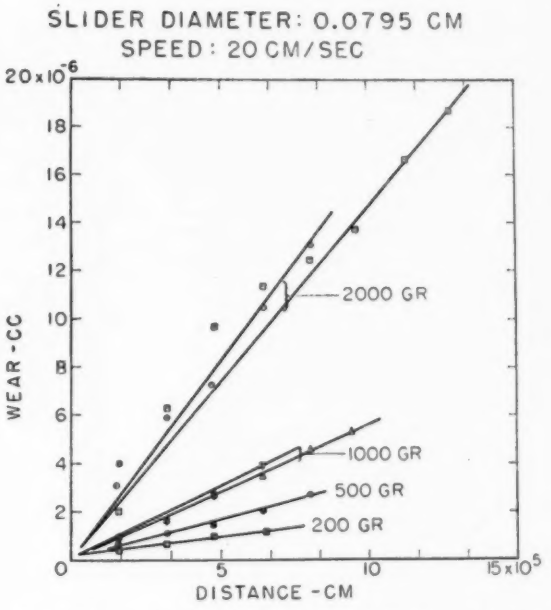


FIG. 6. Wear *vs* distance of travel for cylindrical rider at several comparatively light loads.

It was found that after continued running on the same track on the disk the wear rate would change rather markedly, especially if the applied load had been high enough to cause wholesale scoring of the pin surface. This was probably due to pin material being transferred and welded to the disk to form work-hardened projections which then gouged the pin surface to give cutting wear. When this happened, a new track had to be used. There was sometimes a difference of as much as a factor of two in wear rate between adjacent tracks on the same disk, and between two supposedly identical disks. The rates would sometimes differ by a similar amount. Different batches of hexadecane were also found to give different rates. For this reason in the results below, all data shown on a single plot or compared in any way were taken on a single track. Data in different plots were often from different tracks or disks so that the numerical values are not directly comparable. However, the same qualitative trends and functional relations were always found to hold.

RESULTS

(A) Low Stress

Taking the factors in Eq. (3a) in turn, the dependence of wear on the distance of travel was first investigated. The results are shown in Fig. 6. In this case the two surfaces were carefully run-in so that a steady state con-

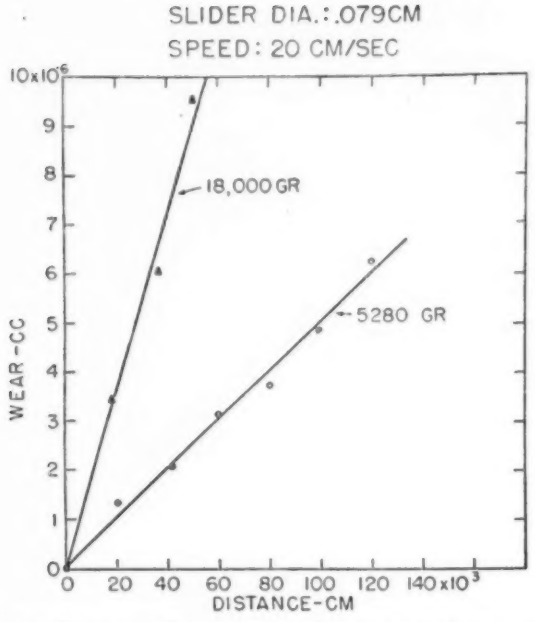


FIG. 7. Same conditions as Fig. 6 at somewhat heavier loads.

dition of operation was reached before any data were taken. The situation during running-in which is of considerable interest in itself will be discussed later. It will be seen from Fig. 6 that the wear is quite accurately linear with distance of travel for a number of different low loads. The same is seen in Fig. 7 to hold also at loads that are somewhat higher but still less than the yield strength of the solid material. The highest load in the figure, 18,000 g, corresponds to about $\frac{2}{3}$ of the yield strength on this diameter rider.

We turn next to the dependence on load. Since the wear *vs* distance is linear for a given load, the ratio V/L, which is the wear rate, is constant and can be conveniently compared against load. This is done in Fig. 8 for cylindrical riders of two different diameters. It is seen

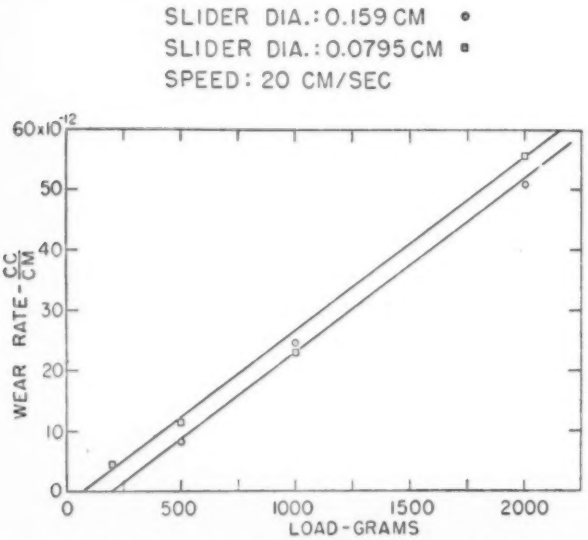


FIG. 8. Wear rate *vs* load for two cylindrical riders.

that the wear rate V/L bears a straight line relation to the load in both cases except that it makes a finite intercept with the load axis. Also the slope of the straight line is seen to be independent of the apparent area of contact, only the intercept changing.

Evidence for this intercept is shown more definitely in Fig. 9 where it is seen that there is a small but finite load for which there is no appreciable wear. (As forementioned, these figures are taken from different sets of experiments in which either the disk or the lubricant or some other condition was changed so that the numerical values of wear rate, etc., are not directly comparable. However, in any given set of runs the qualitative relationships being illustrated by the figures were invariably found to hold.) Figure 8 shows the intercept to increase with size of rider and it is roughly proportional to its apparent area. This load intercept appears to be quite definite and is probably due to hydrody-

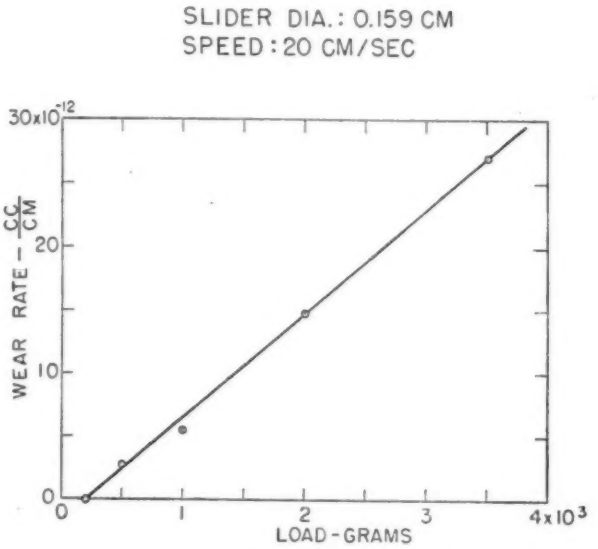


FIG. 9. Wear rate *vs* load for cylindrical rider showing threshold load for wear in presence of a liquid.

namic action of the lubricant used for flooding the disk surface. Hence, this amount of load should be subtracted from the total applied load to give the load that is actually being carried by the solid contact areas. If this is done, then Figs. 8 and 9 show that the wear rate is linear with this net load.

Incidentally, the geometry of the apparatus is such that the hydrodynamic action cannot be caused by the conventional converging wedge of Reynolds which is responsible for most hydrodynamic lubrication. The relative inclination of the two surfaces, if there is any, will be diverging in the direction of motion so that the hydrodynamic pressure in the present case is presumably a result of thermal expansion of the lubricant in passing between parallel or slightly diverging surfaces, as suggested by Fogg.[8] Alternatively, Clark, Woods, and

[8] A. Fogg, Engineering **156**, 138 (1945).

White[9] have recently suggested that such a complete separation of the surfaces can be produced by solidification of a mineral oil under pressure. Another possibility is local waviness of the surfaces producing small Reynolds' wedges. This point is still under investigation.

The wear rate is also linear with load in the higher range of loads as seen from the lower curve in Fig. 10. (On this scale the intercept for this diameter rider would be negligible and the same will be true of all the high load results to be discussed subsequently.) Here again, the highest load corresponds to about $\frac{2}{3}$ the yield strength of the rider material which was the highest to which one could go with the cylindrical shape of rider without failure by plastic deformation.

A linear relation between wear and distance of travel also holds for riders of other shapes such as a truncated cone, as shown in Fig. 11. Here of course the stress

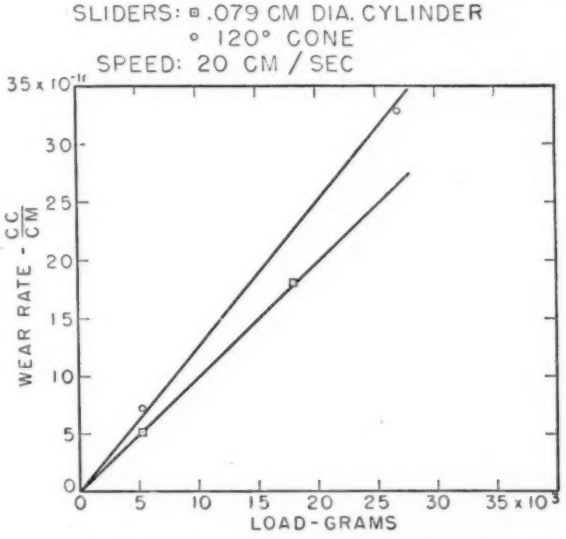


FIG. 10. Wear rate *vs* load for cylindrical and truncated conical riders.

changes as wear progresses because of the change in wear area. This wear rate is also linear with applied load as shown by the second curve in Fig. 10 up to an initial stress of almost 9/10 the yield strength.

From the results presented so far we may conclude that under these conditions, namely,

(a) Well run-in surfaces so that the local areas of true contact are uniformly distributed over the apparent area, i.e., there are no local concentrations of contact,

(b) The average normal stress, likewise quite uniformly distributed over the apparent area, does not exceed the yield strength of the softer material,

then Eq. (3) predicts the correct dependence on distance of travel and on load, provided that in the latter case the hydrodynamic lift or intercept in plots like Figs. 8 or 9, if it is appreciable, is subtracted from the total

[9] Clark, Woods, and White, J. Appl. Phys. **22**, 474 (1951).

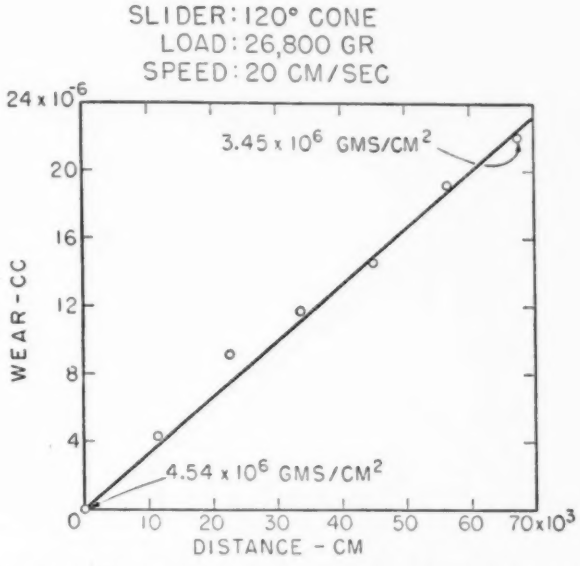


FIG. 11. Wear *vs* distance of travel for truncated conical rider. The values of normal stress at beginning and end of run are also shown.

applied load. At this stage it would appear that regardless of the basic assumptions, Holm's final equation fits

TABLE I.

Rider shape	k/p_m	p_m (bulk)	Knoop hardness
cylinder	0.90×10^{-14} cm²/g	$1.0Y$*	300
120° cone	1.15×10^{-14} cm²/g	$2.6Y$	317

* Y is the yield strength of the rider material.

the facts quite well within the range of conditions outlined above. Results for the higher stress range will be discussed below.

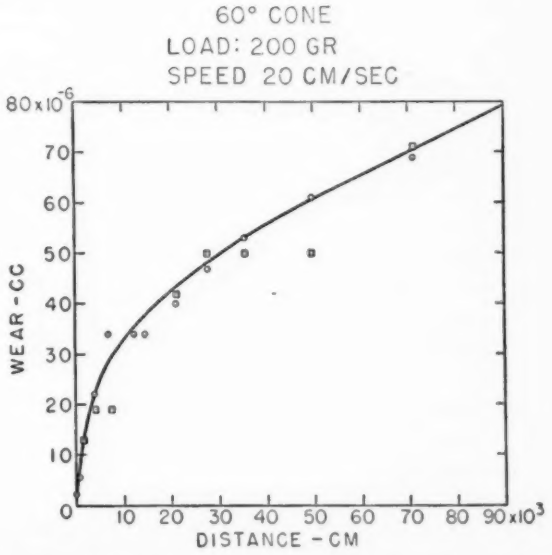


FIG. 12. Wear *vs* distance of travel for conical rider. (Two runs.)

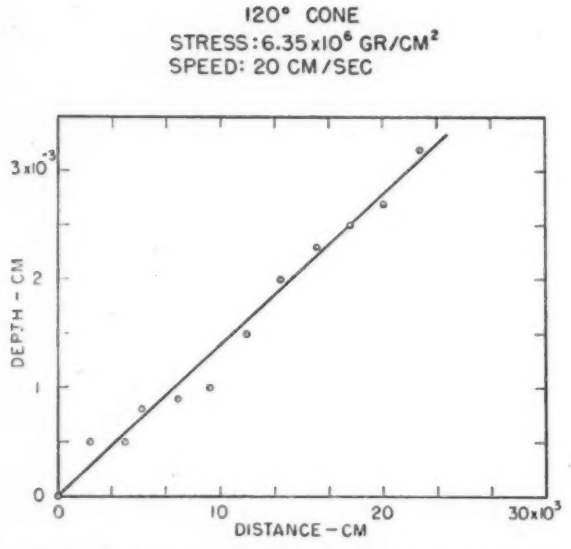


FIG. 13. Depth of wear *vs* distance of travel at constant normal stress (equal to 1.2 times the yield strength) for conical rider.

Incidentally, it might appear that one could also test the dependence on p_m, which appears in Eq. (3), by making cylindrical and conical riders out of the same material and running them under identical conditions. In this case the constant k in Eq. (3) which is determined by the material should certainly be the same in the two cases, and yet, as Bowden and Tabor[10] have shown, the flow pressure, p_m, for these two different shapes is quite different (see Table I). However, these are the values of the flow pressure for the shapes of the bulk rider and not of the local asperities on the rider surface whose yielding determines the true contact area according to Eq. (1). This is borne out by comparison for a cylinder and a 120° cone of the quantity

$$\frac{k}{p_m}=\frac{V}{L\cdot W},$$

which is a rearrangement of Eq. (3a) and is determined from the slope of the straight lines in Fig. 10. The values are given in Table I. It is quite evident that k/p_m does not vary inversely with the bulk p_m. The results simply suggest that the local asperities have the same average shape on the surfaces of both the cylinder and the cone, and consequently, p_m must have the same value. This is borne out by measurement of the superficial (Knoop) hardness of the two surfaces which is seen in Table I to be closely the same. On the other hand hardness measured with the Rockwell $\frac{1}{16}$ in. ball indenter gives values having the ratio of the bulk p_m's. This seems quite reasonable since if one is considering a volume lying in the surface which is small compared to the dimensions of the rider, which would be true of any asperities, then its properties should be independent of the gross shape of the rider. Hence, we must also conclude that in this stress range the wear rate is independent of the gross shape of the rider and that it is not possible to vary p_m by changing the rider shape.

[10] F. P. Bowden and D. Tabor, *Friction and Lubrication of Solids* (Oxford University Press, New York, 1950), Chap. I.

(B) High Stress

In spite of the above agreement with Eq. (3a), consideration of commonly occurring wear curves such as that of Fig. 1 shows that this relation cannot hold generally. In such cases where the load is kept constant, the wear rate is usually high at first and decreases as the rubbing progresses. Results such as this can also be obtained under the conditions of the present experiments by running a conical rider against the disk. The results of two such runs are plotted together in Fig. 12. Here also the wear rate is high at first but later decreases to a constant value such as the truncated cone showed in Fig. 11. However, the friction coefficient was also measured during one of the runs in Fig. 12 by modifying the apparatus to include a force-measuring element and was found to remain constant at 0.12 throughout the run.

One then must ask in the light of Eqs. (3) what is going on to make the wear rate change even though the load is kept constant? There are two obvious differences between the conditions in this run and those of the previous runs with cylindrical riders. The first is that the apparent area and hence the average stress are varying during the run. The second is that during the early part of the run when the apparent area is small, the average stress is considerably higher than any discussed previously, including that of the truncated cone of Fig. 11. To investigate these factors, a run was then made with the conical rider but keeping the average stress constant and at a high value while the apparent area varied during the run. Figure 13 gives the results for the depth of wear *vs* distance which show that now the depth wear rate is constant from the very beginning of the run. The stress used was 1.2 times the yield

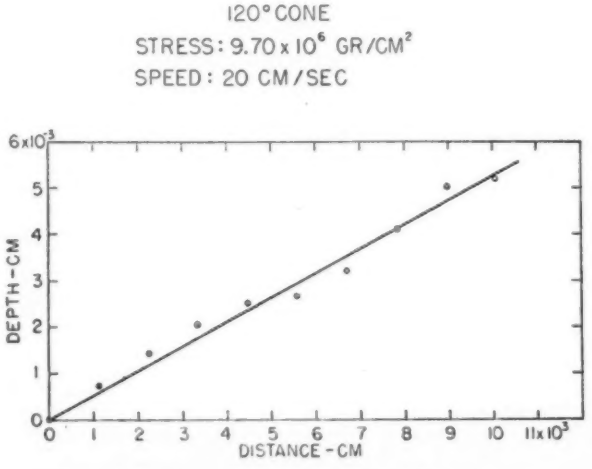


FIG. 14. Same conditions as Fig. 13 at higher stress (equal to 1.8 times the yield strength).

strength Y in tension. It is only possible to impose such a stress on a conical rider of the appropriate cone angle since the flow pressure of the 120° cone used in this run is $2.6Y$. Figure 14 also shows a constant wear rate at a somewhat higher stress of $1.8Y$ on the same conical rider.

Since the depth wear rate h/L is a constant for a constant normal stress, it seems appropriate to plot it against the corresponding stress. However, Eq. (3b) suggests that the quantity $h/L \cdot P$ would be of more interest since in the low stress region where the equation holds this quantity should be a constant. Accordingly, this quantity is plotted in Fig. 15. This plot shows, as expected, that at low stresses the quantity $h/L \cdot P$ is a constant as Eq. (3b) would predict. However, at a rather sharply defined stress the quantity increases very sharply in marked disagreement with the equation. This corresponds to the high wear rate shown in the initial portion of the curve of Fig. 12.

The region of increased $h/L \cdot P$ on the right-hand side of Fig. 15 has several characteristics. First, the wear rate for a given stress immediately becomes 10 to 100 times as great as that predicted by Eq. (3b). Second, the stress at which wear runs could be made was limited by the onset of scoring of the rider surface resulting from transfer of rider material to the disk surface to form welded, work-hardened projections which then gouged the rider surface producing an excessive amount of cutting wear. Third, the same shape of curve as that shown in Fig. 15 was also obtained when the wear runs were made in the absence of a lubricant although the data was more scattered and tended to be obscured by excessive oxidation of the rubbing surfaces. Hence, the sharp break in the curve cannot be a result of breakdown of the liquid lubricant film at some critical stress.

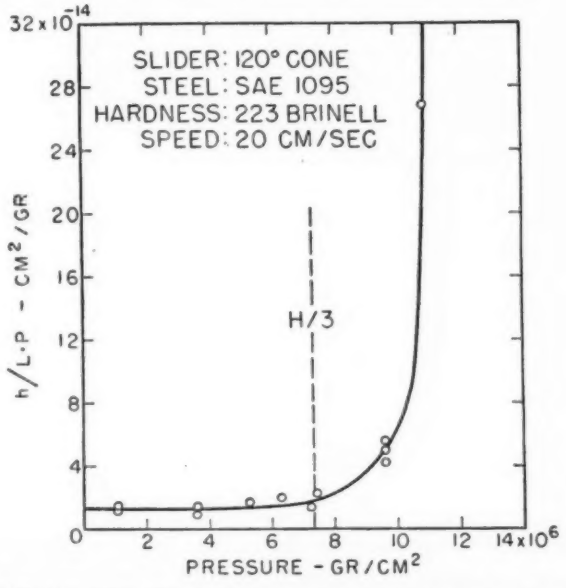


FIG. 15. Summary of data over whole range of stress for conical rider in the form of the ratio of the depth wear rate to the normal stress plotted against the normal stress.

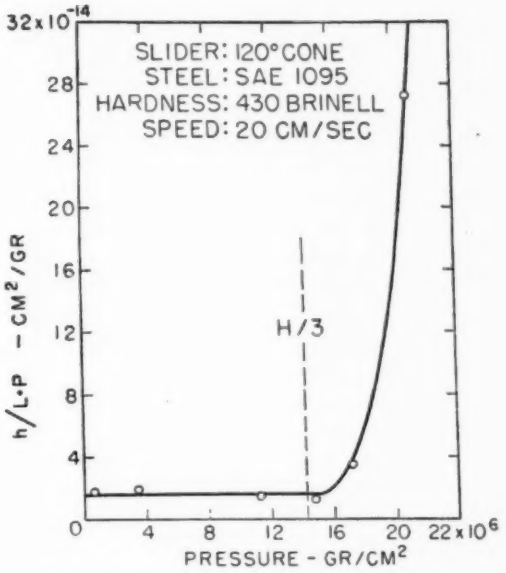


FIG. 16. Similar data to that in Fig. 15 for conical rider of same steel hardened to about twice the hardness.

It is not excluded, of course, that it could be caused by the breakdown of a dry film, either of oxide or other material.

A similar shape of curve is obtained with a conical rider of the same steel (SAE 1095) but hardened to about twice the hardness or 430 Brinell as shown in Fig. 16. It is also characterized by a low stress region where Holm's Eq. (3b) holds, and a higher stress region where the wear rate increases very rapidly with stress to relatively high values.

Considering the low stress regions for the two hardnesses of riders where Eq. (3b) holds, one can calculate relative values of the coefficient k. If it is assumed that p_m in both cases is proportional to the indentation hardness,§ then the relative values of k in the two cases can be calculated by means of Eq. (3b) using the value of $h/L \cdot P$ given by the horizontal portions of the curves in Figs. 15 and 16. The values are given in Table II. It is seen that although the steel is the same in the two cases, k differs by a factor of 2.4. This may come about by either of two causes. First, the process of hardening the steel may very well change the nature of the surfaces since observable changes in microstructure are always found. Second, the shapes of the local asperities on the wear surface of the rider in the two cases may not necessarily be the same, especially since the wear process is

TABLE II.

Rider Hardness	Relative Value of k
223 Brinell	1
430 Brinell	2.4

§ The actual value of p_m must include a factor depending on the shapes of the local asperities. This will lie between approximately 1 and 3(10).

known to yield larger wear particles from softer surfaces. If this were the case, then the p_m's would not be solely in the ratio of the two hardnesses as was assumed in computing the values in Table II.

While the absolute value of k in these cases cannot be computed because we do not know the absolute value of the p_m's, still p_m must be of the order of magnitude of the indentation hardness whence k can be computed to be of the order of 1 to 5×10^{-7}.

Considerable additional data other than those shown here have been obtained and all fall on curves of exactly the shape shown in Figs. 15 and 16, using spherical riders as well as conical riders of different apex angles and both lubricated and dry.

DISCUSSION

Discussion of the data presented above had best be divided into consideration of the two stress regions distinguished by the horizontal and by the steeply rising portion of the curves in Figs. 15 and 16 which, for convenience, we shall call the low stress and high stress regions, respectively.

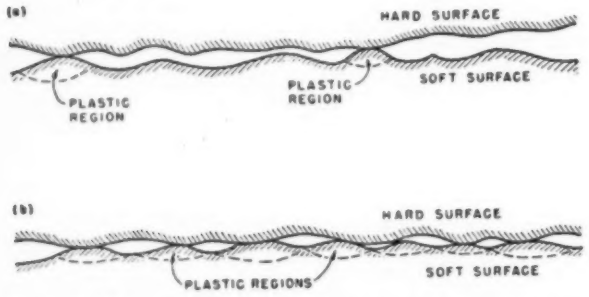


FIG. 17. Sketch of stress distribution in and around local asperities which are in contact. (a) At very low stress. (b) At stress equal to one third of the hardness.

In the low stress region the fact that the quantity $k/L\cdot P$ is a constant means that Eq. (3b) holds. However, as pointed out earlier it seems unlikely that the assumption of material removal atom by atom which Holm made in deriving the equation is correct. Actually, the material is probably removed as relatively large aggregates of atoms. This being the case we must discard the concept of individual atomic encounters. However, if we think in terms of wear particles rather than atoms, it still seems desirable to make use of a concept of probability of removal similar to that suggested by Holm. In this case the constant k in Eqs. (2) and (3) becomes the probability of removing a wear particle per unit distance of travel.

In attempting to estimate the relation between the probability of removing a wear particle and quantities such as the load or true contact area, there is some recently reported information which is pertinent. This is the work of Rabinowicz and Tabor[11] who have studied, by autoradiographic means, the size and distribution of metal particles transferred by welding from one surface to another on sliding. There they found that either dry or in the presence of a poor lubricant, the amount of material transferred was roughly proportional to the load since the main effect of increasing load was to produce a proportional increase in the number of welded junctions or true contact areas rather than in their size. It was also found[12] that the mass of the individual transferred fragment was roughly proportional to the three halves power of its welded area of contact to the other surface indicating that its three dimensions were roughly equal. Furthermore, Fig. 3 suggests that the individual wear particles are approximately the same size. Combining these results suggests the picture that in this low stress region a single wear particle is produced by a single weld, and that increase in stress increases linearly the number of welds, and hence, the number (the probability k remaining constant) of particles being removed. From this Eqs. (2) and (3) follow.

The data obtained showed that the constant k was extremely small, about 1 to 5×10^{-7}. Hence, the probability of removing a wear particle is very small, even with a rather indifferent lubricant, which is indeed fortunate for everyday experience. If Eqs. (2) and (4) are combined there results

$$V=\frac{k}{s}\cdot E, \tag{6}$$

where E is the friction work done in the sliding. If V is multiplied by $Q\rho$ where Q is the latent heat of atomization and ρ the density, then $Q\rho V$ is the energy necessary to remove the material worn off, atom by atom. In other words

$$Q\rho V=\frac{kQ\rho}{s}\cdot E. \tag{7}$$

Whittaker[13] has pointed out that for typical dry wear data the ratio $Q\rho V/E$ of this energy to the friction work is quite small, of the order of 10^{-2} and hence that even in the extreme case that the material were removed atom by atom very little of the friction work goes into producing wear. This is in line with the present finding that k is very small.

Implicit in the above picture is that there is no interaction between the removal processes of individual particles, i.e., a particle once removed is washed completely free from between the rubbing surfaces without producing a second wear particle by scoring action. This seems probable if the true contact area is dispersed into very small local regions widely separated from each other.

Considering next the high stress region where Eq. (3) definitely does not hold the most interesting question posed by the data is the exact nature and explanation

[11] E. Rabinowicz and D. Tabor, Proc. Roy. Soc. (in press).

[12] E. Rabinowicz, Brit. J. Appl. Phys. Supp. No. 1, p. 82 (1951).

[13] E. J. W. Whittaker, Nature **159**, 541 (1947).

of the sharp rise in the quantity $h/L \cdot P$ shown in Figs. 15 and 16. While a detailed explanation is not yet possible at this stage of the investigation, certain characteristics of the high stress region where $h/L \cdot P$ rises so sharply seem definite. In this region the stress is at least one third of the indentation hardness of the rider material and so, according to Eq. (1), the instantaneous true contact area is at least one third or more (depending on the geometrical factor in p_m for the shape of the local asperities) of the nominal area on which wear is taking place. Under these circumstances it seems quite probable that a wear particle removed from one region of the true contact area cannot get away without scoring and producing further wear on other parts of the true contact area downstream. The secondary wear particles in turn produce further wear in a self-accelerating manner and in extreme cases a sort of avalanche of wear material may be formed; there literally is not room between the surfaces for a particle once formed to get out of the way. On this picture increase in stress in this region increases the number of primary wear particles as well as the true contact area; the resultant congestion produces greater numbers of secondary particles. Hence, the wear may be expected to increase much faster than linearly and when the true contact area covers a large enough fraction of the wearing area, a heavy avalanche of wear particles will be produced with complete instantaneous destruction and welding of the surface, as actually observed.

A final question in the explanation of the results concerns the location of the point of transition between the two regions of different wear behavior. In both Figs. 15 and 16 this point is seen to occur at very nearly one third of the indentation hardness of the rider material in question. A possible reason for this is as follows. At low stresses the load is carried on the tops of isolated local asperities as shown in Fig. 17(a). Each of these in turn transmits its share of the load to its respective base where it merges into the bulk metal. The material in the vicinity of the base of each protuberance will begin to flow when the local stress reaches the yield strength of the material which for a completely work-hardened material as the wearing face undoubtedly is, equals approximately one third of the hardness.‖[14] As the load is increased, more individual asperities are brought into contact as discussed above and more of the base material of the surface is brought to this state of stress. Finally, when the applied load reaches one third of the hardness stress as distributed uniformly over the wear area, then all of the base material has been brought to this state of stress as in Fig. 17(b) and further increase in load cannot bring in fresh unstressed surface material.

At this point the stress to produce further plastic flow of the material at the bases of the asperities must increase sharply owing to local interference of adjacent regions and further plastic deformation is shifted preferentially to the peaks of the asperities. This increase in normal stress on the peak material may be expected to do two things. First, it will greatly enlarge each individual contact area already formed, as well as produce new ones, thus resulting in larger wear particles whose size, since they are probably equiaxed, increases with the three-halves power of the average stress on the true contact area and hence with the same power of the load. Second, this increased stress should, according to any of the best accepted criteria of plastic yielding, result in a lowered shear strength of the peaks with consequently greater removal of material. For these reasons a stress of approximately one third of the hardness may be expected to be critical in the wear process.

Turning now to the application of the results of this investigation, it is evident that results like those of Figs. 15 and 16 have a bearing on the running-in of freshly assembled machine parts under a constant load which give curves like that of Fig. 1. At the beginning of running, no matter how carefully the surfaces are made and aligned, they will bear on only an edge or a corner or a local high region. The nominal contact or wear area is thus quite small (far less than the dimensions of the surfaces) and the normal stress on that area is high, well up in the high stress region of Figs. 15 and 16. As a result, the initial depth wear rate is extremely high and the original wear area becomes rapidly enlarged. Care must be taken at this stage to prevent the catastrophic scoring resulting from the self-accelerating process previously described. The run-in may be considered successful if the surfaces are made initially near enough to conformance so that this rapid wear rate will bring their entire areas into mating. When this stage is reached, the stress should have dropped to the low stress region in Fig. 15 where the wear rate is many times less than the initial rate if the surfaces have been properly designed as regards size. In fact this rate is often so small, particularly compared to the initial rate, that it may be negligible and it is often said that wear has ceased when the surfaces have reached this completely run-in condition. This has been variously ascribed to a "glazed" or otherwise specially conditioned surface. While such a surface may be formed in some instances, it seems probable that the successfully run-in surface is simply operating on the left-hand portion of the curve in Figs. 15 and 16.

If the surfaces are not near enough to conformance originally, the initial wear area grows only gradually so that the average stress remains in the high wear rate region for a long time or even indefinitely. Such surfaces appear never to cease wearing.

Still a third possibility is that the wear area grows sufficiently to drop the stress to the low stress, low wear rate region but the wear area is still only a fraction of the designed area. As a result, the surface does not

‖ The yield strength of the bulk material of which the rider in Fig. 15 was made was measured directly in a tensile test and found to be 5.1×10^6 g/cm (72,500 psi) which is appreciably less than one-third of the hardness. The difference, of course, is due to work-hardening of the rubbing surface.

[14] D. Tabor, Proc. Roy. Soc. **A192**, 247 (1948).

function properly, for instance as a gas or oil seal in the case of piston rings. The rings continue to pump oil and yet the wear rate is so low that they will never wear in completely in a normal life-time. Such instances are observed in practice and the purposeful rough surfacing of new piston rings is an insurance that the stress will remain in the high wear rate region until the ring is in contact over most of its face.

If the normal stress is sufficiently low after the run-in, then the hydrodynamic action accounting for the intercept in Figs. 8 and 9 may be sufficient to carry the load, thus separating the surfaces and completely preventing wear. In the examples of Figs. 8 and 9 this hydrodynamic pressure was about 1×10^4 g/cm^2 (142 psi), but its actual magnitude will depend on such factors as the rigidity with which the surfaces are supported, the smoothness of the harder surface, the lubricant viscosity, and the sliding speed.

Thus it can be seen that the results arrived at in this work on the dependence of wear rate on normal stress throw considerable light on many of the observations about running-in.

We should not close this discussion without a strong word of caution about the difficulty of obtaining consistent results even under the carefully controlled conditions of these experiments. All the trends illustrated in this paper could always be duplicated in single runs, but the absolute magnitude of the wear would change greatly from run to run unless the greatest care was exercised to exactly duplicate conditions. Such apparently trivial changes as refinishing the disk surface, using a different batch of the same lubricant, or a change in the relative humidity in the laboratory as mentioned beforehand would change the wear values by 50 or 100 percent. Obviously, these factors have caused changes in the other wear-producing factors such as corrosion, cutting, etc., which are extremely important but which will have to be the subject of a separate investigation once the answers on the basic wear factor—galling—have been obtained. Maintaining accurate alignment of the wearing surfaces once they were run-in was found to be extremely important. For instance, sudden large changes in load would deflect the machine sufficiently in spite of its great rigidity to cause edge loading and consequent increase in wear rate above normal.

CONCLUSIONS

Careful measurements have been made of the wear of rubbing steel surfaces as a function of load, distance of travel, and hardness under conditions which eliminated the effect of all factors except adhesion or galling. The following results were obtained.

(1) At very low pressures averaged over the apparent contact surface the surfaces are completely separated by a hydrodynamic pressure which acts even between worn-in, strictly parallel surfaces.

(2) At somewhat higher average pressures, but still less than one third of the hardness of the softer material, the amount of wear depends linearly on the distance of travel and the load (minus the aforesaid hydrodynamic lift) and is independent of the apparent area of contact.

(3) At average pressures above one third of the hardness, the depth of wear still varies linearly with the distance of travel, is independent of the apparent area of contact for a given average normal stress, but it increases many-fold for small increases in stress. In the limit large scale galling sets in.

(4) These results explain quite well the various characteristics of the running-in process, not only when it is successful but also the observed instances of defective run-in, namely, the case where wear never ceases throughout the useful life of the part and also the case where run-in ceases before the parts are completely mated.

(5) These results suggest that at low stresses a fraction, k, of the local welds which are very small and much isolated each produce a single wear particle of approximately equal size. At high stresses the individual welded areas are large and close together. As a result, larger wear particles result which in turn produce more wear particles before they escape from between the surfaces, in a self-accelerating process.

The authors are greatly indebted to the Chrysler Corporation for its generous support which has made this work possible. They are also indebted to the Chrysler Engineering Laboratories and especially to Dr. Clayton R. Lewis for their cooperation in taking the electron microscope pictures of the wear particles and to Dr. E. Rabinowicz for much helpful discussion.

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 1 JANUARY, 1952

Electromagnetic Resonant Behavior of a Confocal Spheroidal Cavity System in the Microwave Region*

J. C. SIMONS† AND J. C. SLATER
Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts

(Received August 2, 1951)

The resonance of a small spheroidal object in a large spheroidal cavity is investigated. This approximates the case of a thin needle-like antenna in a large cavity. It is shown that this needle, if thin enough, shows marked resonant properties, in that when the cavity is tuned to a resonant frequency defined by the needle, the magnetic field on the surface of the needle is greatly enhanced. This property can be used in a practical way, in measuring surface impedance of the material of which the needle is composed: at resonance most of the loss in the cavity is located at the surface of the needle and depends on the material of which it is composed.

A THIN spheroidal antenna in empty space is known[1] to show resonance effects when its wavelength is approximately a whole number of half-wavelengths; the thinner the antenna and the more closely it approaches a line, the closer the resonant lengths are to whole numbers of half-wavelengths. Also the Q of the resonance becomes infinite as the antenna becomes infinitely thin. These facts suggest that a thin spheroidal antenna in a resonant cavity might show similar results. If so, it would have a practical value for certain microwave measurements. It is often desired to investigate the surface impedance of a sample of material; that is, its contribution to the Q of a cavity containing it. If the whole cavity is made of the material in question, the calculation of the relation between Q and the surface impedance is simple. If, however, it is only convenient to introduce a sample of one material into a cavity of another, both materials will make contributions to the Q of the composite cavity. It is then obviously desirable to concentrate the tangential magnetic field, and related surface current, on the surface of the sample as far as possible, so that it will contribute most of the losses, and a comparatively small part will come from the remaining wall of the cavity. The facts regarding high Q antennas in empty space suggest that if the sample were made in the form of such an antenna, the larger part of the loss would be concentrated on the sample. Pippard[2] has used resonant samples in his studies of the surface impedance of conductors and semiconductors at very low temperatures and microwave frequencies.

To investigate this problem, a natural method is to solve the case of resonance of a confocal spheroidal cavity, the inner spheroid representing the sample antenna, the outer spheroid the cavity. This case can be handled analytically, and it is possible to compute both the resonant frequencies as functions of the various parameters, and the ratio of the contributions of inner and outer conductors to Q. Solutions of this problem have been obtained for a number of values of the parameters.[3] We shall not present the detailed results in the present note. The solutions elucidate the qualitative nature of the result, however, and that is simple enough so that it seems worth while to describe it briefly.

Let us start by fixing the distance between the foci of the spheroids. The infinitely thin spheroid will then be a line or needle, whose length is the interfocal distance. Special resonance properties will be shown for those frequencies for which this infinitely thin antenna is $l+1$ half-wavelengths long (so that l measures the number of nodes of the magnetic field in the length of the antenna). Let us now set up a standing electromagnetic wave of one of these resonant frequencies, with the appropriate number of nodes along the length of the needle, and satisfying the proper boundary conditions on the surface of the antenna. This standing wave will have an infinite number of confocal spheroidal surfaces surrounding the antenna, distant approximately a half-wavelength from each other, on which the tangential electric field will be zero, so that any one of these surfaces could be replaced by a metallic cavity wall without interfering with the radiation pattern. The larger spheroids, with many wavelengths inside them, will of course approach spheres. Such a cavity, then, is so chosen that the frequency of the resonant modes corresponding to a given l value with it will equal the frequency of the corresponding mode of the free antenna. We may call such a cavity a standard cavity.

The solution we have just set up is not the only type consistent with the resonant frequencies. At any such resonant frequency, we have not only the solution just described, which satisfies boundary conditions on an infinitely thin spheroid; we have also an independent solution, related to the first solution as the sine is to the

* This work has been supported in part by the Signal Corps, the Air Materiel Command, and the ONR.

† Now at Westinghouse Atomic Power Division.

[1] J. A. Stratton and L. J. Chu, J. Appl. Phys. **12**, 230 (1941); S. A. Schelkunoff, Proc. Inst. Radio Engrs. **29**, 493 (1941); and later work. For discussion, see J. C. Slater, *Microwave Transmission* (1942), Chapter V.

[2] A. B. Pippard, Proc. Roy. Soc. (London) **A191**, 370–415 (1947).

[3] J. C. Simons, "Electromagnetic resonant behavior of a confocal spheroidal cavity system in the microwave region," M.I.T. thesis (1950).

cosine at large distances, and having a singularity at the origin of such variety that it cannot satisfy suitable boundary conditions at a needle-shaped antenna. We can then make any linear combination of these two solutions, and each such linear combination will have a family of confocal spheroids on which the tangential component of electric field is zero. We could make two such spheroids, of which one may be small, the other large, into conductors, and thus set up a resonant solution of this type, and a standard cavity, for an inner spheroid of any arbitrary eccentricity.

We may now take such a standard cavity and perturb it by either increasing or decreasing the size of the outer wall, keeping it always a confocal spheroid and leaving the inner spheroid unchanged. As we do this, we can follow the behavior of each of the resonant modes of the cavity, and each such mode will have its frequency tuned, but without change of its fundamental nodal structure. As the cavity is shrunk the frequency will increase, since we still have the same number of nodes between the antenna and the wall, and this distance decreases, bringing a decrease of wavelength, or increase of frequency; similarly, increase of the size of the cavity decreases the frequency. The effect of these changes on the wave pattern may be described qualitatively. As the cavity shrinks and the frequency increases, the field tends to decrease near the antenna, and at the same time the region of high field near the antenna shrinks down toward the center of the antenna, so as to preserve the correct wavelength. If the cavity is greater than the standard size the variation is of the opposite sort, the field being stronger nearer the antenna.

The essential point for our purpose, however, is not this variation of field strength, but one particular quantity to be derived from it, the integral of H^2, where H is the magnetic field, over the inner spheroid, divided by the corresponding integral over the outer spheroid; that is, the ratio of losses on inner and outer spheroid, assuming them to be made of the same material. When we make the calculations we find that the loss on the inner spheroid is small for frequencies lower than the resonant frequency, increases as the frequency increases, goes through a maximum, and then decreases again at still higher frequencies, the losses on inner and outer spheroids becoming equal at infinite frequency. The thinner the inner spheroid, the smaller the frequency range over which the loss of the inner spheroid is greater than half its maximum value, or the greater the Q as measured by the breadth of the resonance peak; also the thinner the spheroid, the higher the maximum value of loss on the inner spheroid. For very thin spheroids the maximum loss on the antenna comes almost exactly at the resonance frequency, corresponding to the standard cavity. For thicker spheroidal antennas, however, the maximum shifts to lower frequencies. Both these features of the behavior, the dependence of resonant frequency and of the breadth of the resonant peak on the thickness of the antenna, are qualitatively similar to the situation found for antennas in free space.

The calculations which have been carried out have been only exploratory, directed toward elucidating the qualitative nature of the results, as described in the present note; in particular, the methods used have not allowed the evaluation of results for frequencies much greater than resonance and are not adapted to very thin antennas and high Q. More work would be necessary to get quantitative value of the Q of the resonance (as measured by maximum loss of the inner spheroid, or half-width of the resonance) and the resonant frequency as functions of the eccentricity or thickness of the inner spheroid. The results clearly confirm, however, the supposition that by using a very thin antenna it is possible to concentrate most of the field on the antenna; to do this, the outer cavity must be tuned to a frequency closely agreeing with the free-space resonant frequency of the antenna. Since this result is of such a simple qualitative nature it is to be anticipated that a similar resonant antenna put inside some other shape of external cavity will show similar results as the cavity is tuned to the resonant frequency of the antenna. We may plausibly generalize somewhat further, and suppose that other shapes of resonant objects, as for instance objects shaped like small Hertz oscillators, put inside of ordinary cavities, would show similar properties: as the cavity is tuned to a frequency which would be a resonant frequency of the oscillator if it were in free space, the field near the oscillator will build up greatly, so that most of the loss in the cavity will be found on the surface of the oscillating antenna, very little on the surface of the cavity.

The Structure of Sputtered Silver Films*

C. E. Ellst and G. D. Scott
Department of Physcis, University of Toronto, Canada
(Received July 10, 1951)

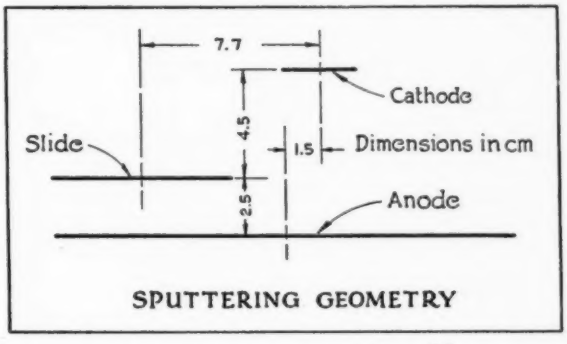
The electron microscope has been used to study the structure of sputtered silver films. When compared with evaporated silver films which are prepared at the same rate, the sputtered films are observed to have a more continuous structure. When compared with evaporated silver films which are produced at much faster rates, the sputtered films exhibit a crossover phenomenon in which the thinner films are the more continuous and the thicker films less continuous than evaporated films of the same thickness. These properties are in agreement with resistivity measurements. A possible explanation for the results is proposed.

INTRODUCTION

OVER the past twenty years sputtered metal films have lost much of their technological importance in favor of evaporated films. This may be chiefly because of the marked advances in the techniques of vacuum evaporation. However, as a consequence, most of the recent research on thin metal films has dealt with evaporated films. In this connection the electron microscope has been used extensively and has provided useful information regarding the factors effecting the structure and hence the properties of the films. In particular, the existence of an aggregated structure,[1] the way in which the structure depends on rate of formation,[2] and the correlation of observed structure with electrical resistivity[3] and with optical properties,[1] have been established. On the other hand the literature apparently contains only one reference to the application of the electron microscope in studying sputtered films.[4] It therefore seemed desirable to examine sputtered films with the electron microscope and to correlate their structure and properties in the light of the information already available for evaporated films.

Several earlier attempts to compare the properties of sputtered and evaporated films have been reported.[5–8] These comparisons have included evaporated films formed both in high vacuum and in the residual atmosphere used for sputtering. Apart from certain anomalies appearing at the geometrical center of the sputtered deposit the two types of films, i.e., sputtered and evaporated, have the same general appearance, and their electrical resistivities are qualitatively similar. Faust[9] has made a critical survey of the optical properties of silver films as given in individual reports and concludes that the two types of film are similar in this respect also. But such a comparison unless carried out by experiment in the same laboratory is of necessity very qualitative in nature. By electron diffraction it has been clearly established that both types of films are polycrystalline with the same crystallographic structure as the bulk metal.[10]

Silver was used throughout the work described here because of its advantages in sputtering[11] and because data on evaporated silver films were readily available.



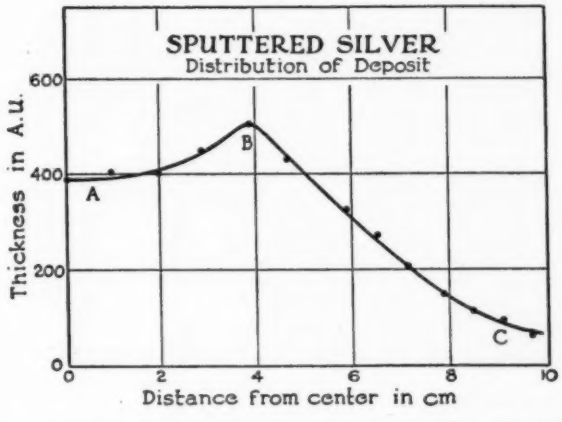


FIG. 1. The geometrical arrangement for sputtering and the radial distribution of the deposit.

* Published with the permission of the Director General, Department of Mines and Technical Surveys, Ottawa, Canada.
† Now at the Physical Metallurgy Research Laboratories, Ottawa.
[1] R. G. Picard and C. S. Duffendack, J. Appl. Phys. **14**, 291 (1943).
[2] R. S. Sennett and G. D. Scott, J. Opt. Soc. Am. **40**, 203 (1950).
[3] T. A. McLauchlan, Ph. D. thesis, University of Toronto, 1950.
[4] Physical Laboratory of Louvain (to be published): see Lambeir, van Itterbeek, and van den Berg, Physica **16**, 907 (1950).
[5] W. Reinders and L. Hamburger, K. Akad. Amsterdam Proc. **19**, 858 (1917).
[6] H. Kahler, Phys. Rev. **18**, 210 (1921).
[7] L. R. Ingersoll and L. O. Sordhal, Phys. Rev. **32**, 649 (1928).
[8] L. R. Ingersoll and J. D. Hanawalt, Phys. Rev. **34**, 972 (1929).
[9] R. C. Faust, Phil. Mag. **41**, 1238 (1950).
[10] G. I. Finch and H. Wilman, Ergeb. Exakt. Naturwiss. **16**, 353 (1937).
[11] E. N. Andrade and J. G. Martindale, Trans. Roy. Soc. (London) **A235**, 69 (1935).

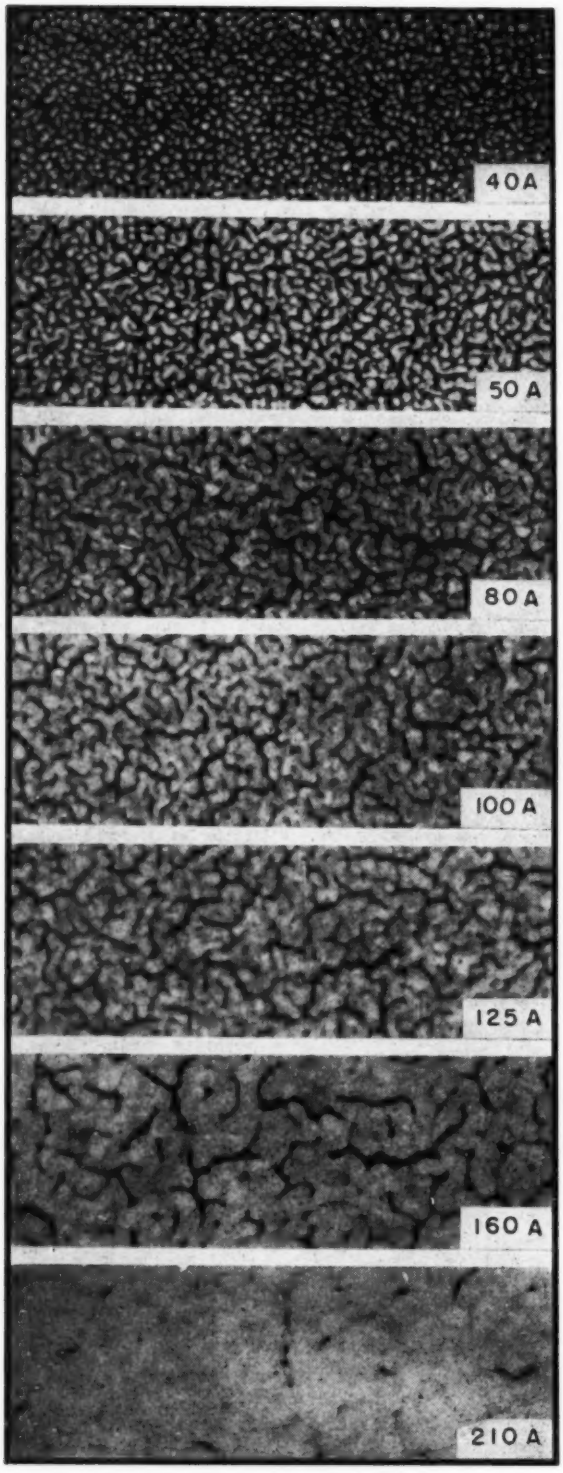


FIG. 2. Electron micrographs of sputtered silver films at 75,000×. Average thicknesses are indicated in Angstrom units. Note that the contrast has been reversed in the reproductions to make the silver particles appear light.

EXPERIMENTAL

The films were sputtered in an atmosphere of air at 30μ, using voltages of 1100 and 2000 volts dc. The geometry of the electrodes is shown in Fig. 1; the anode was a sheet of aluminium, the cathode a plane coil of pure silver wire. The resistivity measurements were taken from films formed on a glass microscope slide. Formvar mounts placed at the edge of the slide were used for the electron microscope observations. The position of the slide avoided the region of anomolous deposits, and by providing an even gradation of film thickness along its length, as shown by the graph of Fig. 1, allowed the study of a range of thicknesses from one sputtering run. It has been shown that the variation in angle of incidence of the incoming particles along the length of the slide and also the change in substrate from glass to the Formvar used for the microscope mounts have no noticeable effect on the structure or the properties of the films.[2]

The substrate temperatures in this investigation were measured with a chromel-alumel thermocouple pressed against the slide surface at its center. For all films described here the temperature varied only between room temperature and 10°C above. Experimental work on the aggregation of silver films because of substrate temperature indicated that this small variation would not be significant. Since an increased sputtering voltage resulted in a slightly increased substrate temperature any dependence of the structure on the voltage alone was not easily determined, but tests showed that the structure was not sensitive to the range of voltage used here.

The microscope slides were cleaned by washing in Orvus solution, rinsed in distilled water and polished with lens paper. The cathode was cleaned before each run by sputtering it for about five minutes. Resistance measurements were made as soon as possible after the sputtering, using the method outlined by McLauchlan.[3] All micrographs were taken on the day the film was formed. The thickness measurements were made by multiple beam interferometry.[12]

The structure of the sputtered silver films at different thicknesses for a deposition time of twenty minutes is shown by the micrographs in Fig. 2. These may be compared with the micrographs of evaporated silver films given by Sennett and Scott[2] (see Fig. 3(a) of this reference). The comparison shows that for the same rate of deposition the structure of the sputtered films is considerably more continuous at all values of the thickness. This same result is shown by the resistivity curves of Fig. 3, where the higher resistivity of the evaporated films indicates a greater aggregation in the structure.

Some films were evaporated in air at the pressure of the sputtering atmosphere. The resistivity curve for such a film is given also in Fig. 3. For thicknesses below 300A these films covered more of the substrate than did the vacuum evaporated but the aggregates did not link together to the same extent as those of the sputtered.

[12] Scott, McLauchlan, and Sennett, J. Appl. Phys. **21**, 843 (1950).

The resistivity curves of Fig. 4 indicate that when the 20-minute sputtered films are compared with the 2-second evaporated, a crossover occurs in the region of 120A. Thus the sputtered film should be more continuous at thicknesses below this crossover, the evaporated at thicknesses above. A comparison of Fig. 2 with Fig. 2(a) of Sennett and Scott[2] shows that the electron micrographs confirm this effect.

DISCUSSION

Though the literature contains much speculation on the nature of the sputtered particles the mechanism by which they leave the cathode has not yet been clearly demonstrated.[13] One of the more generally accepted theories suggests that an evaporation occurs from local areas of the cathode which have been heated to a high temperature by impact of the positive ions.[14] If this is the case the initial velocity of the atoms in the sputtering process will be approximately the same as in the evaporation process. However, in sputtering, the velocity of the particles will be reduced by collisions with the molecules of the gas so that their velocity as they approach the substrate will be considerably less than in vacuum evaporation. On such a basis then, it would be expected that the vacuum evaporated films would show more aggregation because of the possibility of increased migration of the atoms over the substrate. This is as observed, but the results of the evaporation at sputtering pressure, where the atom velocities will also be reduced, show that there must be other factors in addition which influence the film structure.

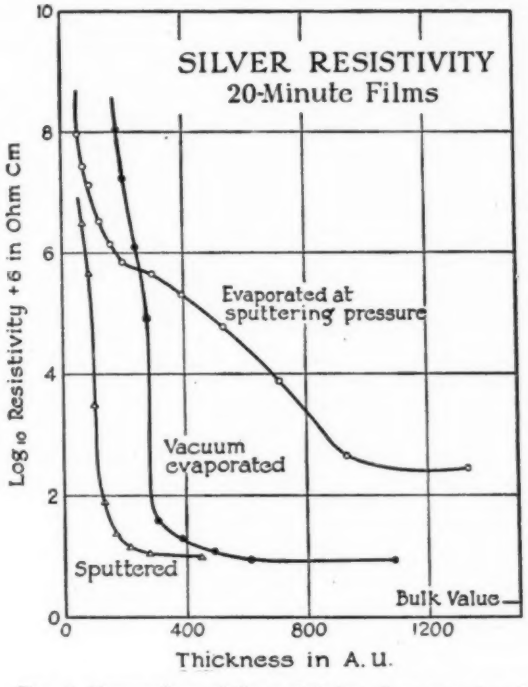


FIG. 3. Comparison of the resistivity of sputtered and evaporated silver films.

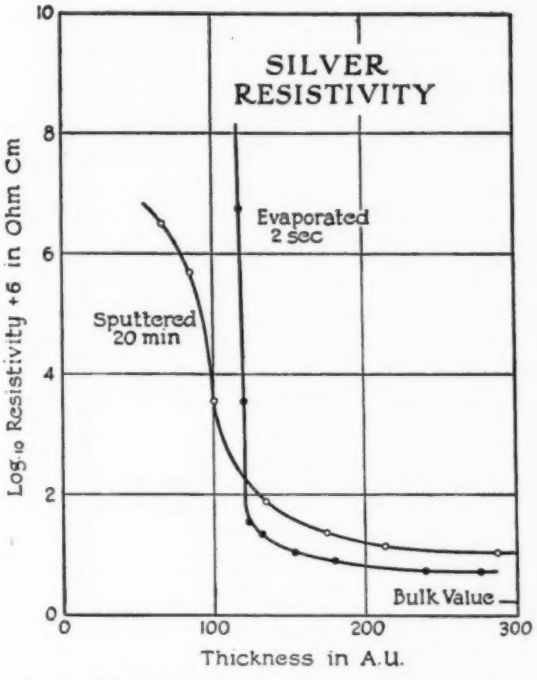


FIG. 4. The resistivity curves with thickness for 20-minute sputtered and 2-second evaporated films.

The connection between the structure of thin metal films and the phenomenon of "critical density" of condensation of metal atoms on a substrate has been pointed out by Levinstein.[15] From experimental observations that a deposit of metal on glass can only be be secured under certain conditions of surface temperature and density of incoming atoms, Knusden[16] formulated the concept of the probability, ϵ, of an atom being permanently adsorbed on striking a substrate surface. For any one metal and surface condition there exists some critical density at which a sharp change of ϵ occurs, this density separating ranges of rapid and slow condensation. Frenkel[17] has worked out a quantitative theory which predicts that the effect should occur. Ditchburn[18] attempted to observe this "critical density" effect in sputtered films, but he concluded that if the effect was present at all it would be at densities far below that for evaporated particles.

To explain his observations Ditchburn postulated that positive ions (or some other particle) in the discharge act as a cleansing agent to remove a layer of adsorbed air from the substrate surface over the small area at which the ion struck. On these cleaned areas the metal atoms have a much better opportunity of being permanently captured by the surface. By using the concept of van der Waals adsorption of the metal

[13] A. Aron, Ann. Phys. (Paris) **1**, 361 (1946).
[14] E. Blechsmidt and A. von Hipple, Ann. Phys. (Leipzig) **86**, 1006 (1922).
[15] H. Levinstein, J. Appl. Phys. **20**, 306 (1949).
[16] M. Knusden, Ann. Phys. (Leipzig) **50**, 472 (1916).
[17] J. Frenkel, *Z.* Physik **26**, 117 (1924).
[18] R. W. Ditchburn, Proc. Roy. Soc. (London) **A141**, 169 (1933).

atoms on the surface, it is not difficult to show that the heat of adsorption of a metal atom on gas covered glass (as will be the case generally at room temperature even in high vacuum) is considerably less than it is on a clean glass surface. This lowered heat of adsorption results in greater mobility of the atoms over the surface and hence tends to produce films with more aggregation.

The Frenkel theory assumes that an atom striking the substrate is not immediately completely adsorbed, but moves on the surface as in a two-dimensional gas. After a certain time the atom leaves the surface, unless it collides with another atom, in which case the pair have a much longer life on the surface. These doublets can form the nuclei around which the film is built. It follows then that at equal rates the sputtered films should be more continuous than the evaporated, for the sputtered

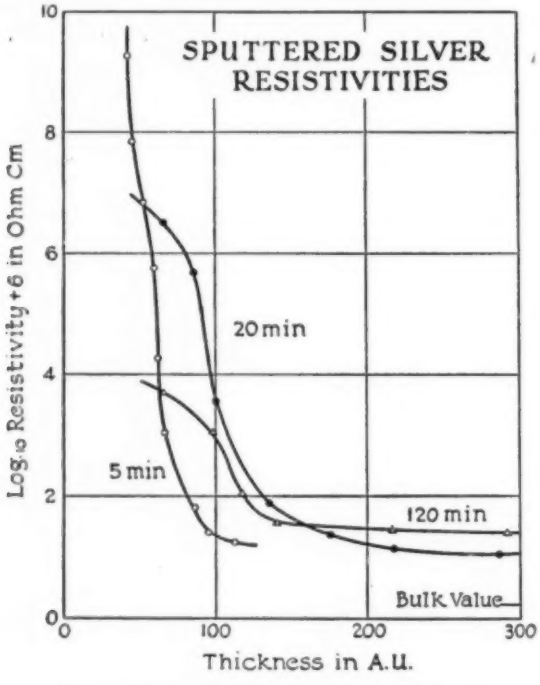


FIG. 5. The effect of rate of deposition on the resistivity of sputtered silver films.

atoms will have all the opportunity of the evaporated atoms to form these doublets, and in addition there will be the nuclei provided by the atoms captured on the cleaned areas. In the air atmosphere of this experiment such cleaning may be effected by positive or negative ions.

The possibility of the crossover in the properties of the 2-second evaporated and the 20-minute sputtered films can be explained on the above basis by first considering the influence of rate of deposition. It has been shown that in general a higher rate will produce a more continuous film.[2] There are probably two effects occurring in this phenomenon: firstly, at the higher rates a greater proportion of the incoming atoms will go to form the initial doublets, and secondly, in rapid evaporations there will not be sufficient time for plastic deformation of the aggregates to take place so that they will grow as flat platelets rather than becoming nearly spherical as in the early stages of slow deposition. The crossover will occur when the factors of rate and cleaning take on the same relative importance. For thicknesses below 120A the nuclei which occur in the 20-minute sputtered film on the cleaned areas are more important than the higher rate of formation in the 2-second evaporated film; at thicknesses above 120A the rapid condensation of the doublets and the shaping of the aggregates become important and make the evaporated film more continuous.

It had been planned at the start of this investigation to look for the influence of rate on sputtered films, but the formation of very fast (say 2-second) films, which would be desirable for this purpose, presents considerable experimental difficulty. Micrographs of 5- and 20-minute films gave no clear indication of any major structural differences. However, the resistivity measurements gave evidence of some effect because of rate of deposition. Figure 5 shows the resistivity curves of sputtered films produced in three different times. It is seen that these curves act in a consistent manner. For the thinner films the resistivity decreases with decreasing rate. For the thicker films the resistivity decreases with increasing rate as would be expected from the results on evaporated films. These effects can be explained as follows. When the absolute rate at which the silver comes to the substrate is high (as for the thicker films) the structure is governed by the rate considerations already discussed, but when the absolute rate is low (as for the thinner films) the increased relative cleaning makes the slower films the more continuous.

It can be noted from Fig. 3 that there is a considerable difference between the resistivities of the two types of evaporated films at the greater thicknesses. This is assumed caused by the presence of a larger amount of occluded air in the films evaporated at the sputtering pressure.

CONCLUSIONS

In conclusion it may be noted that the differences in structure between sputtered silver films as compared to evaporated films are probably caused, (1) by the slower speeds of the sputtered atoms as they approach the substrate because of collisions with the residual gas, and (2) by the cleaning action of the discharge on the substrate. Both these effects will tend to make the sputtered films more continuous, and such will be the case for very thin films. However the rate of deposition also affects the structure, higher rates giving greater continuity, and since it is experimentally possible to obtain much higher rates of deposition by evaporation, more continuous films in the thicker ranges can be made by the evaporation process. Thus for most optical and electrical applications of thin metal films, where the required thicknesses are usually greater than about 100A, evaporation rather than sputtering is to be preferred.

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 1 JANUARY, 1952

Simulation of Trajectories of Charged Particles in Magnetic Fields

GEORGE H. VINEYARD
Department of Physics, University of Missouri, Columbia, Missouri
(Received July 23, 1951)

A ball rolling on a rotating surface is shown to simulate the motion of a charged particle in a magnetic field. The theory is given for the case of a warped surface undergoing arbitrary rotation about a fixed axis and translation perpendicular thereto, while the system from which the ball is observed partakes of similar but independent motion. With approximations based on not too large departure of the surface from flatness, the following cases can be simulated: (a) The magnetic field is homogeneous and constant. The electric field is perpendicular to the magnetic field, and irrotational, but otherwise arbitrarily spatially dependent, and arbitrarily time dependent within certain limits. (b) The magnetic field is homogeneous but arbitrarily time dependent. The electric field is perpendicular to the magnetic field and may have a variety of space and time dependences, including a part which encircles the axis and has just the right magnitude to produce acceleration in a circular orbit as in the betatron.

Results of rudimentary experiments are presented which indicate that the method is capable of good accuracy.

INTRODUCTION

COMPUTATION of charged particle orbits in electric fields alone can be carried out expeditiously by several well-known mechanical simulators. The presence of a magnetic field introduces a complication, because of the peculiar velocity dependence of the magnetic force, and relatively little work has been reported on the simulation problem in combined electric and magnetic fields.[1]

We wish to point out that a sphere rolling on a table which is rotating about an axis perpendicular to its surface is acted upon by a force analogous to the force of a magnetic field on a moving charge, namely, a force proportional to the speed of the sphere and perpendicular to its path. By the use of this principle, a mechanical device consisting of a ball rolling on a rotating surface can be made to duplicate the orbits of charged particles moving in certain kinds of electric and magnetic fields. A second way in which a "magnetic" effect can be obtained is by viewing a moving body from a rotating frame of reference, since the Coriolis force likewise has this form. By tipping or warping the surface on which the ball rolls a class of forces dependent on position and time can be introduced, thus duplicating the effect of an electric field. Arbitrary translatory motion of either the table or the frame of reference will introduce the effect of additional forces spatially independent but arbitrarily time dependent. Finally, the combination of a rotating and arbitrarily translating warped surface on which a ball rolls, with a rotating and arbitrarily translating frame of reference from which the ball is observed, will be seen to broaden still further the class of orbits which the ball can be made to simulate. It is the theory of such a device which we present in this paper.

[1] E. Brüche and A. Recknagel [Z. tech Physik **17**, 126 (1936)] show that the case of a pure magnetic field with axial symmetry can, with transformation, be treated by a gravitational model. A. Rose [J. Appl. Phys. **11**, 711 (1940)] describes a device in which the reaction of a gyroscope is used to simulate the magnetic force.

The acceleration, **a**, of a particle of velocity **v**, charge e, mass M, moving in a magnetic field **B** and electric field **E** is given by

$$\mathbf{a}=(e/M)[\mathbf{v}\times\mathbf{B}+\mathbf{E}]. \tag{1}$$

We ask that the rolling ball, as seen from a certain frame of reference, possess an acceleration given by a function of position, velocity, and time equal to the right side of Eq. (1), or derivable from the right side of this equation by a change of length and time scales. We restrict our considerations to the component of the acceleration in one plane, and, hence, our method will be directly applicable only to plane orbits. The magnetic field simulated will have to be perpendicular to this plane and spatially homogeneous; it may be time varying. The electric field simulated may vary with both time and position under certain rather loose restrictions which will be stated later. Some approximations and assumptions, which should be valid under a variety of experimentally attainable conditions, are necessary. These are: (a) The rolling is perfect, that is, there is no slipping and no dissipative friction. (b) The slope of the surface is everywhere small compared with unity. (c) The radius of the ball is small compared to the smallest radius of curvature of the surface. (d) $v'\ll\rho_0\Omega$, where v' is the speed of the ball relative to the surface, ρ_0 is the least radius of curvature of the surface, and Ω is the angular velocity of the surface. (e) The component of angular velocity of the ball parallel to the angular velocity of the table is initially zero, and the orbit is short enough that this component does not become large. For simplicity we also suppose that the rotations of the surface and of the frame of reference are about the same fixed direction, and that the displacements of the surface and the frame of reference are both in a plane perpendicular to the axis of rotation.

THEORETICAL FOUNDATIONS

A sphere rolls on a rigid surface, S, which may be undergoing rotation about a fixed direction and also

translation in a plane perpendicular to this direction.

We adopt the following symbols (see Fig. 1):

s = radius of the sphere.
s_g = radius of gyration of the sphere.
m = mass of the sphere.
O, x, y, z = an origin and set of axes in an inertial system.
O', x', y', z' = an origin and set of axes attached to S. The z' axis is parallel to the z axis, and O' lies in the x, y plane.
$\boldsymbol{\Omega} = \mathbf{k}\Omega$ = angular velocity of S. $\mathbf{k}$ is a unit vector along the z axis. $\Omega = |\boldsymbol{\Omega}|$.
$\mathbf{q}$ = vector displacement of O' from O.
$\mathbf{r}'$ = vector from O' to the center of the sphere.
$\mathbf{r}$ = vector from O to the center of the sphere. $\mathbf{r} = \mathbf{r}' + \mathbf{q}$.
$\mathbf{v} = d\mathbf{r}/dt$.
$\mathbf{a} = d\mathbf{v}/dt$.
$\mathbf{n}$ = unit vector normal to S at the point of contact with the sphere.
$\mathbf{P}$ = the force exerted by S on the sphere.
$\mathbf{G}$ = the acceleration of gravity. $\mathbf{G}$ is not necessarily parallel to $\boldsymbol{\Omega}$.

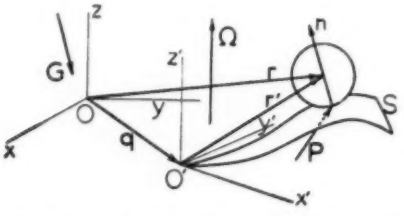


FIG. 1. Ball on moving surface. Fixed and moving coordinate systems.

$\mathbf{H}$ = any additional forces on the sphere besides $\mathbf{P}$ and gravity. It is assumed that $\mathbf{H}$ exerts no torque about the center of the sphere.
$\boldsymbol{\alpha} = d\boldsymbol{\omega}/dt$.
$\boldsymbol{\omega}$ = angular velocity of the sphere.
$\boldsymbol{\omega}'$ = angular velocity of the sphere relative to the system O'. $\boldsymbol{\omega}' = \boldsymbol{\omega} - \boldsymbol{\Omega}$.
$\varphi(x, y, t) = z$; $\varphi'(x', y') = z'$ — equations of S in system O and system O', respectively.
Subscript xy on any vector denotes the (vector) component of that vector in the x, y plane.
Subscript z on any vector denotes the (vector) component of that vector parallel to the z axis.
Subscript t denotes the time derivative.
d/dt' denotes time derivative of a vector relative to the system O'. $d/dt' = d/dt - \boldsymbol{\Omega}\times$.[2]

The equations of motion of the sphere are

$$m\mathbf{a} = \mathbf{P} + m\mathbf{G} + \mathbf{H}, \tag{2}$$

in translation, and

$$ms_g^2\boldsymbol{\alpha} = \mathbf{P}\times\mathbf{n}s \tag{3}$$

in rotation. The condition that rolling occurs without slipping is met by requiring that, relative to the system O', the velocity of the point on the sphere in contact with S be zero:

$$(d\mathbf{r}'/dt') + s\mathbf{n}\times\boldsymbol{\omega}' = 0. \tag{4}$$

Recalling the relation between time derivatives of vectors in the two systems,[2] and the relations between $\mathbf{r}$ and $\mathbf{r}'$, $\boldsymbol{\omega}$ and $\boldsymbol{\omega}'$, Eq. (4) may be rewritten

$$\mathbf{v} - \mathbf{q}_t - \boldsymbol{\Omega}\times\mathbf{r} + \boldsymbol{\Omega}\times\mathbf{q} + s\mathbf{n}\times\boldsymbol{\omega} - s\mathbf{n}\times\boldsymbol{\Omega} = 0. \tag{5}$$

After crossing Eq. (2) into $\mathbf{n}$, the unknown force $\mathbf{P}$ may be eliminated between Eqs. (2) and (3) to give

$$m\mathbf{a}\times\mathbf{n} = (m\mathbf{G} + \mathbf{H})\times\mathbf{n} + (m/s)s_g^2\boldsymbol{\alpha}. \tag{6}$$

By splitting each of the vectors into two parts, one parallel and one perpendicular to the z axis, both Eq. (5) and Eq. (6) may be written as two separate equations:

$$\mathbf{v}_z + s\mathbf{n}_{xy}\times\boldsymbol{\omega}_{xy} = 0, \tag{7}$$

$$\mathbf{v}_{xy} - \mathbf{q}_t - \boldsymbol{\Omega}\times\mathbf{r}_{xy} + \boldsymbol{\Omega}\times\mathbf{q} + s(\mathbf{n}_z\times\boldsymbol{\omega}_{xy} + \mathbf{n}_{xy}\times\boldsymbol{\omega}_z) - s\mathbf{n}_{xy}\times\boldsymbol{\Omega} = 0, \tag{8}$$

from Eq. (5), and

$$m\mathbf{a}_{xy}\times\mathbf{n}_{xy} = (m\mathbf{G}_{xy} + \mathbf{H}_{xy})\times\mathbf{n}_{xy} + (m/s)s_g^2\boldsymbol{\alpha}_z, \tag{9}$$

$$m(\mathbf{a}_z\times\mathbf{n}_{xy} + \mathbf{a}_{xy}\times\mathbf{n}_z) = (m\mathbf{G}_z + \mathbf{H}_z)\times\mathbf{n}_{xy} + (m\mathbf{G}_{xy} + \mathbf{H}_{xy})\times\mathbf{n}_z + (m/s)s_g^2\boldsymbol{\alpha}_{xy}, \tag{10}$$

from Eq. (6).

From the equation for S, we find

$$\mathbf{n} = \boldsymbol{\nabla}(z-\varphi)/|\boldsymbol{\nabla}(z-\varphi)| = [\mathbf{k} - \mathbf{i}(\partial\varphi/\partial x) - \mathbf{j}(\partial\varphi/\partial y)] \times [1 + (\partial\varphi/\partial x)^2 + (\partial\varphi/\partial y)^2]^{-\frac{1}{2}}.$$

To first order in the slopes, this gives

$$\mathbf{n}_z = \mathbf{k}, \quad \mathbf{n}_{xy} = -\mathbf{i}(\partial\varphi/\partial x) - \mathbf{j}(\partial\varphi/\partial y) = -\boldsymbol{\nabla}\varphi. \tag{11}$$

These quantities should be computed at the point $(\mathbf{r} - s\mathbf{n})_{xy}$, but because of the smallness of s compared to the radii of curvature of S and the smallness of the slope of S, we may understand by $\boldsymbol{\nabla}\varphi$ and $\boldsymbol{\nabla}\varphi'$, hereafter, the value computed at $\mathbf{r}_{xy}$. $\mathbf{n}_{xy}$ is seen to be a quantity small to first order in the slope.

From Eq. (9) it now appears that $\boldsymbol{\alpha}_z$ is also small to first order in the slope, and, hence, if $\boldsymbol{\omega}_z$ is zero at the beginning of an orbit, it will be a small quantity throughout the orbit. The term in Eq. (8) involving $\boldsymbol{\omega}_z$ may now be dropped because it is of second order. Then Eq. (8) can be solved for $\boldsymbol{\omega}_{xy}$, with the use of Eq. (11) and expansion of triple vector products, giving

$$\boldsymbol{\omega}_{xy} = s^{-1}[\mathbf{k}\times\mathbf{v}_{xy} - \mathbf{k}\times\mathbf{q}_t + \Omega\mathbf{r}_{xy} - \Omega\mathbf{q}] - \Omega\mathbf{n}_{xy},$$

and, taking the time derivative,

$$\boldsymbol{\alpha}_{xy} = s^{-1}[\mathbf{k}\times\mathbf{a}_{xy} - \mathbf{k}\times\mathbf{q}_{tt} + \Omega(\mathbf{v}_{xy} - \mathbf{q}_t) + \Omega_t(\mathbf{r}_{xy} - \mathbf{q})] - \Omega_t\mathbf{n}_{xy} - \Omega d\mathbf{n}_{xy}/dt. \tag{12}$$

[2] H. Goldstein, *Classical Mechanics* (Addison-Wesley Press, Inc., Cambridge, Massachusetts, 1950), Sec. 4-8.

The last two terms may be dropped, the term containing $\mathbf{n}_{xy}$ being small to first order in the slope, and $d\mathbf{n}_{xy}/dt$ being of the order of $\mathbf{v}'/\rho+\boldsymbol{\Omega}\times\mathbf{n}_{xy}$, where ρ is the radius of curvature of the intersection with S of a normal plane tangent to the orbit. Taking the time derivative of Eq. (7) we demonstrate that $\mathbf{a}_z$ is small to first order, thus justifying the omission of the $\mathbf{a}_z$ term from Eq. (10). We now eliminate $\boldsymbol{\alpha}_{xy}$ from this latter equation, with the aid of Eq. (12), solve for $\mathbf{a}_{xy}$ and find

$$\mathbf{a}_{xy}=s_g^2(s^2+s_j^2)^{-1}[\mathbf{q}_{tt}+\boldsymbol{\Omega}\times(\mathbf{v}_{xy}-\mathbf{q}_t)+\boldsymbol{\Omega}_t\times(\mathbf{r}_{xy}-\mathbf{q})]$$
$$+s^2(s^2+s_g^2)^{-1}[\mathbf{G}_{xy}+m^{-1}\mathbf{H}_{xy}-(G_z+m^{-1}H_z)\mathbf{n}_{xy}]. \quad (13)$$

For the especially simple case of a flat, horizontal surface rotating at constant speed about a vertical axis, and with $\mathbf{H}=0$, Eq. (13) reduces to $\mathbf{a}_{xy}=s_g^2(s^2+s_g^2)^{-1}\boldsymbol{\Omega}\times\mathbf{v}_{xy}$, which reveals the primary effect of interest, the pseudo-magnetic force supplied by the rotating table. In this case, incidentally, all of our approximations become exact.

We now introduce a second moving system, a reference system from which the motion of the ball may be viewed (see Fig. 2). The following additional symbols will be employed:

O'', x'', y'', z'' = the origin and axes of the second moving reference system. The z'' axis is parallel to the z axis, and O'' lies in the x, y plane.

$\mathbf{p}$ = the vector displacement of O'' from O.

$\boldsymbol{\Gamma}=\mathbf{k}\Gamma$ = the angular velocity of the system O''. $\Gamma=|\boldsymbol{\Gamma}|$.

Position, velocity, and acceleration of the ball relative to the system O'' are denoted by double primes on appropriate symbols. In particular $\mathbf{r}''=\mathbf{r}-\mathbf{p}$.

$c=s_g^2/(s^2+s_g^2)$.

$\mathbf{E}^*$, $\mathbf{B}^*$ = effective electric and magnetic fields, respectively, such that a particle of charge e and mass M would have the same motion in them as the ball in our model.

Other conventions are analogous to those already employed.

From the relation[2] $d/dt=d/dt''+\boldsymbol{\Gamma}\times$, we find

$$\mathbf{v}=\mathbf{v}''+\mathbf{p}_t+\boldsymbol{\Gamma}\times\mathbf{r}'',$$

and

$$\mathbf{a}=\mathbf{a}''+2\boldsymbol{\Gamma}\times\mathbf{v}''+\boldsymbol{\Gamma}\times(\boldsymbol{\Gamma}\times\mathbf{r}'')+\boldsymbol{\Gamma}_t\times\mathbf{r}''+\mathbf{p}_{tt}.$$

Substituting these in Eq. (13) we find

$$\mathbf{a}_{xy}''=\mathbf{v}_{xy}''\times(2\boldsymbol{\Gamma}-c\boldsymbol{\Omega})+\Gamma(\Gamma-c\Omega)\mathbf{r}_{xy}''$$
$$+(c\boldsymbol{\Omega}_t-\boldsymbol{\Gamma}_t)\times\mathbf{r}_{xy}''+c\boldsymbol{\Omega}_t\times(\mathbf{p}-\mathbf{q})$$
$$+c\boldsymbol{\Omega}\times(\mathbf{p}_t-\mathbf{q}_t)+c\mathbf{q}_{tt}-\mathbf{p}_{tt}$$
$$+(1-c)[\mathbf{G}_{xy}+m^{-1}\mathbf{H}_{xy}-(G_z+m^{-1}H_z)\mathbf{n}_{xy}]. \quad (14)$$

This is the desired equation of motion of the center of the ball as viewed from the moving system O''.

The parameter c describes the pertinent mechanical properties of the ball. c always lies between 0 and 2/5, and for a uniform solid sphere $c=2/7$. The effective magnetic field, $\mathbf{B}^*$, equals $(M/e)(2\boldsymbol{\Gamma}-c\boldsymbol{\Omega})$, and the effective electric field, $\mathbf{E}^*$, equals (M/e) multiplied by all the terms on the right side of Eq. (14) except the first one. The first contribution to $\mathbf{E}^*$ is a centrifugal force (which is inward if $c\Omega>\Gamma$), and the second is a tangential force caused by the angular accelerations. The next three terms in $\mathbf{E}^*$ depend on the relative displacement of the two origins, and the final term is the contribution of gravity plus any additional body forces on the sphere, $\mathbf{H}$. Several arrangements utilizing forces of this type can be imagined; thus, the genuine magnetic attraction of a permanent magnet for a steel ball might be helpfully employed. However, these involve technological complications beyond our present scope, and henceforth we assume $\mathbf{H}=0$. The gravitational term $\mathbf{G}_{xy}$ arises when the axes of rotation are not vertical, and contributes to $\mathbf{E}^*$ a vector constant in magnitude, independent of position, and rotating with an angular velocity $-\boldsymbol{\Gamma}$. For many applications the axes may be set vertically, and then the entire gravitational contribution to $\mathbf{E}^*$ becomes

$$(M/e)(1-c)G\mathbf{n}_{xy}=-(M/e)(1-c)G\nabla\varphi'(x',y'), \quad (15)$$

where the gradient is evaluated at the point x'', y'', related to x', y' by a simple transformation involving the time.

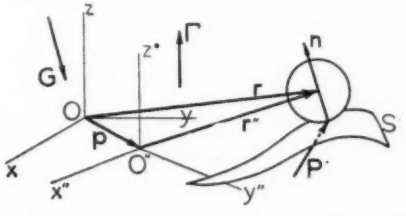


FIG. 2. Ball on moving surface. Fixed coordinate system and second moving coordinate system.

APPLICATIONS

Many special cases of possible interest can be deduced from Eq. (14), and we do not attempt an exhaustive treatment. The principal possibilities may be outlined as follows:

I. B^* Is Constant

Hold Ω and Γ constant. Then,

$$\mathbf{B}^*=(M/e)(2\boldsymbol{\Gamma}-c\boldsymbol{\Omega}) \quad (16)$$

$$\mathbf{E}^*=(M/e)[\Gamma(\Gamma-c\Omega)\mathbf{r}_{xy}''+c\boldsymbol{\Omega}\times(\mathbf{p}_t-\mathbf{q}_t)+(c\mathbf{q}_{tt}-\mathbf{p}_{tt})$$
$$+(1-c)(\mathbf{G}_{xy}-G_z\mathbf{n}_{xy})]. \quad (17)$$

A. $\mathbf{E}^*$ *Is Independent of Time.*

Choose $\mathbf{p}=\mathbf{q}=0$, axes vertical, and $\boldsymbol{\Gamma}=\boldsymbol{\Omega}$. Then,

$$\mathbf{B}^*=(M/e)(2-c)\boldsymbol{\Omega},$$

$$\mathbf{E}^*=(M/e)(1-c)[\Omega^2\mathbf{r}_{xy}''-G\nabla\varphi'(x',y')].$$

We see that by appropriate choice of φ' any *irrotational*

FIG. 3. Simulation of an orbit in a constant magnetic field with no electric field.

$\mathbf{E}^*$ can be simulated. A second arrangement in the case that $\mathbf{E}^*$ has rotational symmetry might be advantageous; choose $\Gamma=c\Omega$, or $\Gamma=0$. This eliminates the centrifugal force term, simplifying the demands on φ'.

B. $\mathbf{E}^*$ *Varies with Time.*

1. $\mathbf{E}^*$ *Is Spatially Homogeneous.*—Choose $\Gamma=0$ or $\Gamma=c\Omega$, $\varphi'=0$, and axes vertical. Then

$$\mathbf{E}^*=(M/e)[c\boldsymbol{\Omega}\times(\mathbf{p}_t-\mathbf{q}_t)+(c\mathbf{q}_{tt}-\mathbf{p}_{tt})].$$

A variety of ways of choosing time varying $\mathbf{p}$ and $\mathbf{q}$ exists to give $\mathbf{E}^*$ any desired time dependence.

2. $\mathbf{E}^*$ *Is Not Spatially Homogeneous.*—In the most general case, all the terms of Eq. (17) will be retained, and a wide variety of time dependent electric fields can be simulated. The limitations are these: $\mathbf{E}^*$ is irrotational and is always of the following form—

$$c_1\mathbf{r}_{xy}''+c_2\nabla\varphi'(x',y')+\mathbf{g}(t), \qquad (18)$$

where c_1 and c_2 are constants, $\mathbf{g}(t)$ is an arbitrary (vector) function of time, φ' is an arbitrary function, the ∇ is computed at x'', y'', t, and x'' and y'' are related to x', y' by an arbitrarily time dependent displacement and a rotation at a constant rate. This case should have interesting applications to the orbit problem in the magnetron.

II. $\mathbf{B}^*$ Is Time Dependent

Ω and Γ are allowed to vary. Restricting our attention to rigid surfaces, the centrifugal force term will impart an undesirable time dependence to $\mathbf{E}^*$ unless Ω and Γ are properly related. We consider only the cases in which this term remains zero, requiring that either $\Gamma=c\Omega$ or $\Gamma=0$.

A. $\Gamma=c\Omega$.

Then $\mathbf{B}^*=(M/e)c\boldsymbol{\Omega}$, and can be given an arbitrary time dependence. $\mathbf{E}^*$ will be of the form $c_2\nabla\varphi'(x',y')+\mathbf{g}(t)$, where the symbols are as in Eq. (18), except that the rotation of axes linking x', y' and x'', y'' is not at a constant rate.

B. $\Gamma=0$.

Then $\mathbf{B}^*=-(M/e)c\boldsymbol{\Omega}$. The third term of Eq. (14) now contributes to $\mathbf{E}^*$ a term of the form $(M/e)c\boldsymbol{\Omega}_t\times\mathbf{r}_{xy}''$, say $\mathbf{E}_1^*$. Then $\mathbf{E}_1^*=-\mathbf{B}_t^*\times\mathbf{r}_{xy}''$, and this is precisely the relation between electric and magnetic fields at the orbit which prevails in the betatron. If, than, a ball is placed on a horizontal flat table, initially at rest, and the table is set into rotation, the ball will simultaneously commence circulating around the axis of the table, gaining velocity at just the right rate to remain at constant distance from the axis, in good imitation of the acceleration of electrons in the betatron. This phenomenon can easily be observed. The $\mathbf{p}$ and $\mathbf{q}$ terms and the gravitational terms can then readily be employed to simulate a variety of perturbing electric fields besides the induction field, $\mathbf{E}_1^*$. The discussion of possibilities parallels that of case I, B.

EXPERIMENTAL CHECKS

A rudimentary experimental check on the theory has been made with a very simple apparatus. A 16-inch disk was cut from a sheet of plate glass and fastened to a turntable of the sort commonly used in physics lecture demonstrations. The table was spun by hand or allowed to coast freely (under which condition the angular deceleration was about 0.01 rev/sec^2). A $\frac{3}{4}$-inch ball bearing served as the rolling sphere.

Figures 3–5 are multiflash photographs of some of the resulting motions as observed from a fixed system. The interval between flashes was 0.01 second in every case. The image of the light source, reflected by the ball, shows as a series of small, bright dots. Also visible are

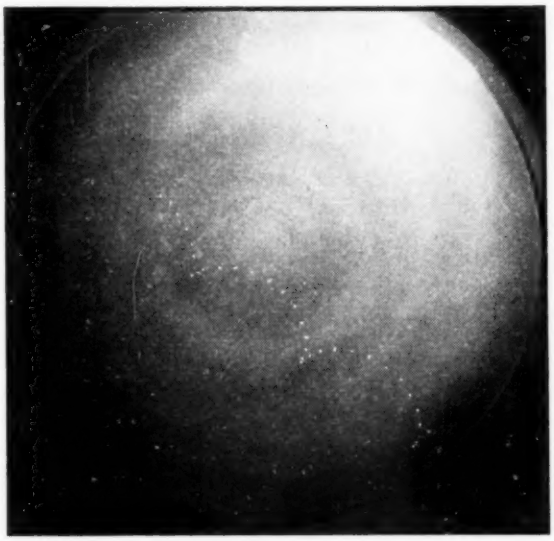


FIG. 4. Simulation of an orbit in a constant magnetic field with constant perpendicular electric field.

satellite images caused by reflections in the glass table top, and these should be disregarded.

Figure 3 shows a case in which the table was spinning freely with its surface horizontal. This is one of the cases under heading I, A, of the previous section, and simulates a constant magnetic field with zero electric field. As expected, the ball pursued a circular orbit at constant speed. This orbit had a diameter of 2.3 inches, and the angular velocity of the table was about 0.51 rev/sec. The fact that the period of the orbit was almost exactly an integral number of tenths of a second was fortuitous. With care, orbits of two or three revolutions, nicely circular and closely superimposed could be obtained.

Figure 4 shows a case in which the table was tipped about $2\frac{1}{2}°$ from the horizontal and left spinning freely (at a rate of 2.36 rev/sec). This simulates the case of a constant magnetic field perpendicular to the table and a constant electric field tangent to the table. The ball was released from rest (after being allowed to come up to rotational speed), and described a cycloidal orbit. The point of release was in the upper left corner of the photograph and does not show. The direction of maximum slope of the surface was accurately perpendicular to a line through the cusps of the cycloid.

In Fig. 5 the table was level and initially at rest, with the ball at rest near the lower left corner of the picture. The table was given a clockwise angular acceleration, and the ball began moving around the center of the table with synchronous acceleration. This illustrates the betatron action, case II, B. Although the orbit is closely circular here, there is a departure from perfection in that the center of the orbit is displaced by about 10 percent of its diameter from the center of the table. It is believed that this was caused by a small unidirectional tipping of the table, as it was accelerated by hand and there was some looseness in the bearing.

Apart from this displacement from center of the betatron orbit, the fidelity of these motions to the theory is quite satisfactory. Thus, the orbit in Fig. 3 is circular to within 2 percent of its diameter, and the relation between observed orbital angular velocity and mean angular velocity of the table checks the theory to about 1 percent. Similarly the dimensions and velocities in the orbit of Fig. 4 are correct to within about 2 percent. The glass surface used departed from flatness by 0.007 inch at several points, and was also somewhat scratched. As already mentioned, the bearing of the turntable had appreciable looseness, and there was some

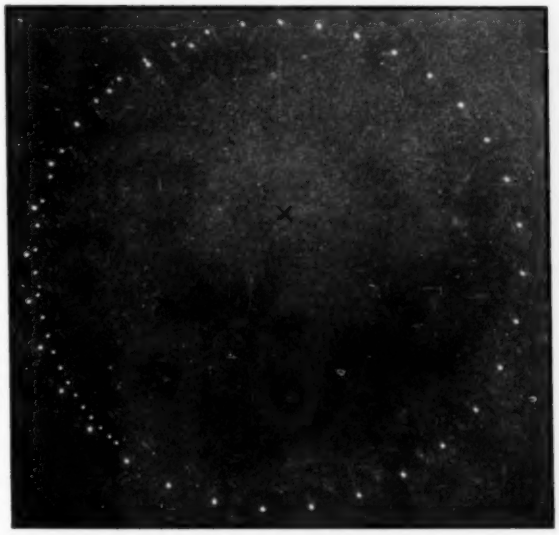


FIG. 5. Simulation of an orbit in a magnetic field which increases with time and an induced electric field, as in the betatron. Diameter of orbit is 5 inches. Center of table indicated by X.

deceleration of the table when it was supposed to be at constant speed. Thus, these experiments were quite crude, and it would seem that a considerable improvement in conditions could be attained without great effort. Consequently we have reason to believe that the method is capable of achieving quite respectable precision, and that the assumptions made concerning the perfection of the rolling can be closely realized.

No quantitative experiments employing warped surfaces or moving systems of reference have been made. The realization of a moving reference system is basically very easy, of course; the camera is merely translated and rotated in a suitable way during the exposure. For direct observation the rotation can be achieved by viewing the motion through a Dove prism that is being turned at an angular velocity $\frac{1}{2}\Gamma$.[3] Finally, the effect of displacement can be introduced by other arrangements of prisms or mirrors.

The writer is indebted to R. M. Sutton for a discussion which led to this work.

[3] This arrangement has been suggested by Klebba and Stommel [Am. J. Phys. **19**, 247 (1951)] for demonstrating the Coriolis force. They also suggest using, in this demonstration, a ball rolling on a rotating surface. However they give no quantitative analysis, and seem to ignore the effect of a rotating reference frame in altering the angular equations of motion of the ball. The use of a Dove prism was earlier suggested to the writer by R. M. Sutton.

The Supersaturation of Nitrogen in a Hypersonic Wind Tunnel*

I. Faro, T. R. Small, and F. K. Hill
Applied Physics Laboratory, The Johns Hopkins University, Silver Spring, Maryland
(Received July 16, 1951)

Tests in a small hypersonic wind tunnel have determined critical temperature and pressure conditions for the condensation of pure dry nitrogen and have indicated that in the range of pressure (0.5–2.8 mm Hg) and temperature (52°–64°R) obtained in the tests, there is a supersaturation of about 30 Fahrenheit degrees.

I. INTRODUCTION

In the production of flow at Mach numbers of the order of 10, condensation of impurities in the working fluid, and even of the working fluid itself presents many difficulties. The large area ratio between test section and nozzle throat necessary for high Mach number flow and the short nozzle length needed to minimize viscous effects, require such a rapid expansion of the gas that in general the usual vapor-pressure equilibrium conditions are not attained. The degree to which supersaturation may occur, and its effects on the flow properties, constitute one of the basic problems in current hypersonic research. It is well known that condensation may be initiated either by particles of "impurity" in the gas and/or by what Oswatitsch[1] calls molecular germs of the gas itself. The study of the condensation of nitrogen in air is complicated by the presence of water vapor and carbon dioxide which form solid particles at temperatures above the critical temperature for nitrogen itself and hence may prevent the supersaturation of nitrogen from occurring. Several theoretical studies (Charyk and Lees,[2] Bogdonoff and Lees,[3] Stever and Rathbun[4]) have been concerned with (a) the condensation of nitrogen in air and/or (b) the condensation of nitrogen as a pure fluid. Some experimental investigation has also been made of the former by Stever and Rathbun,[4] but to our knowledge no work has been published on the latter. The experiments to be described in this paper were carried out with the purest nitrogen that could be obtained commercially (see Sec. II) in order that the properties of this one gas could be studied at temperatures in the range of 52°–64°R and at pressures from 0.5 to 3 mm Hg. The supply conditions were varied from ambient temperature to 1900°R at storage pressures greater than 1000 psi. The nominal Mach number of the flow at the testing station in the nozzle was 9.

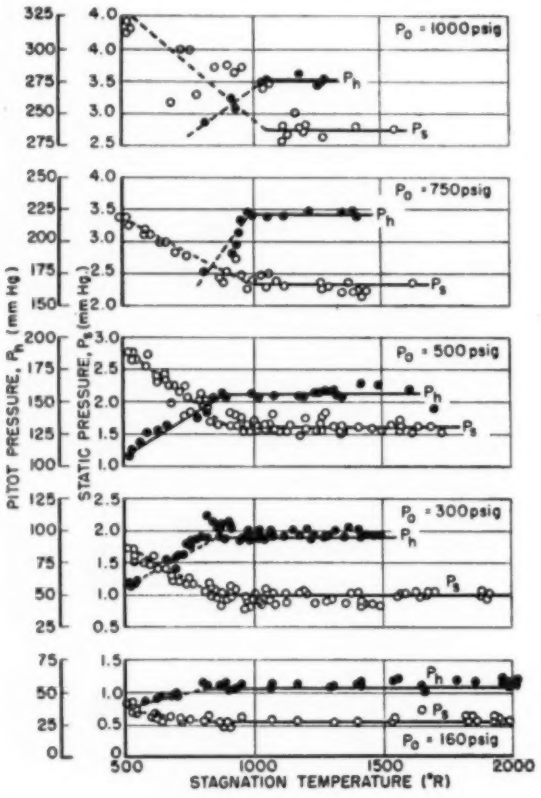


FIG. 1. Pitot-head, p_h, and static pressure, p_s, measurements as a function of stagnation temperature. Full lines represent average pressures (above critical temperature), dashed lines represent trends only.

II. TEST PROCEDURE

The hypersonic nozzle which was used consisted essentially of a simple conical section, which has been described elsewhere.[5] A high pressure—high temperature settling chamber was designed for use with heated nitrogen. The supply pressure was obtained from bottled nitrogen at 2000 psi. The gas analysis as supplied by the Southern Oxygen Company reported 99.8 percent nitrogen and inert gases, 0.2 percent oxygen, no trace of carbon dioxide, and 0.01 to 0.007 mg per liter of water vapor at N.T.P. The gas was heated by passing through six feet of closely coiled stainless steel tubing which served as the resistive element in a low voltage—high current electrical circuit; the mass flow was controlled by a pressure regulating valve. The stagnation temperature, T_0, was regulated by the electrical input and was

* Based on work done at the Applied Physics Laboratory, Johns Hopkins University under Contract NOrd 7386 with U. S. Bureau of Ordnance.

[1] K. Oswatitsch, V D.I. 86, No. 47/48, 702 (1942).
[2] J. V. Charyk and L. Lees, Report No. 127, Princeton Univ. A.E.L. (March, 1948).
[3] S. M. Bogdonoff and L. Lees, Report No. 146, Princeton Univ. A.E.L. (May, 1949).
[4] H. G. Stever and K. C. Rathbun, M.I.T., A.E.D. report, (NACA) (October, 1950).
[5] Faro, Small, and Hill, J. Appl. Phys. 22, 220 (1951).

TABLE I.[a]

1	2	3	4	5	6	7	8	
			Pressure ratios and Mach numbers derived therefrom assuming isentropic flow with $\gamma=1.4$					
Nominal stagnation pressure p_0 psig	Measured pitot-head pressure p_h mm Hg	Measured static pressure (wall and probe) p_s mm Hg	p_h/p_s / M_1	p_h/p_0 / M_2	p_s/p_0 / M_3	Critical stagnation temperature T_0 crit. °R	Test section temperature T_s °R	Degrees of supersaturation Fahrenheit
160	54.43 (26)	0.57 (27)	95.0 / 8.56	6.38×10^{-3} / 8.52	6.72×10^{-5} / 8.53	800	52	31
300	95.34 (34)	0.99 (58)	96.0 / 8.61	5.95×10^{-3} / 8.65	6.20×10^{-5} / 8.64	855	55	31
500	156.3 (16)	1.61 (67)	96.9 / 8.66	5.89×10^{-3} / 8.67	6.08×10^{-5} / 8.67	910	57	30
750	221.6 (8)	2.33 (21)	96.5 / 8.64	5.72×10^{-3} / 8.71	5.92×10^{-5} / 8.70	985	61	28
1000	276.7 (5)	2.76 (10)	100.14 / 8.83	5.34×10^{-3} / 8.82	5.33×10^{-5} / 8.83	1060	64	26

[a] Parentheses give number of test points.

measured at a point in the settling chamber close to the nozzle throat by an iron-constantin or a chromel-alumel thermocouple and recorded on a Brown temperature potentiometer.

The stagnation pressure, p_0, was measured in the settling chamber at a point close to the nozzle throat. The static pressure was measured at a point on the tunnel wall nine inches downstream of the nozzle throat. The usual pitot-static probe was not used, as the extent of the effect of the blunt nose and its normal shock upon the static pressure immediately downstream of the nose is not fully known at high Mach numbers. Pitot head, p_h, and free stream static pressure, p_s, measurements were made independently, the respective measurements being made at a point on the tunnel axis directly above the wall pressure port. The pitot head probe was made of hypodermic tubing sealed in $\frac{1}{8}$-inch stainless steel tubing. The static pressure probe had four ports leading to a common duct. They were spaced at 90° to each other and were ten diameters from the 20° conical tip. The static pressures were measured with diaphragm gages and McCleod gages; the pitot pressures were measured with diaphragm and Zimmerli gages and with absolute pressure manometers. Five series of runs were made with stagnation pressures of 160, 300, 500, 750, and 1000 psig. In each series the stagnation temperature was varied from run to run throughout the range from ambient temperature to about 1900°R. The maximum attainable temperature decreased with increasing stagnation pressure.

III. RESULTS

The measured pressures are shown as a function of stagnation temperature in Fig. 1. The static pressures include both those measured on the tunnel axis by means of the probe and those measured on the tunnel wall. It may be seen that the pressure changes quite rapidly with increasing temperature until a critical temperature is reached beyond which the pressure remains virtually constant. In the 300-psi-case where a great many runs were made, there is evidence of a peak in the pitot pressure in the region of the critical temperature. At this temperature there is not only a wider scatter in the values of the static pressure, but in general a higher pressure is measured by the probe than is measured at the wall.

The measured pressures are given in Table I, Columns 2 and 3. The pressure ratios p_h/p_s, p_h/p_0, and p_s/p_0 are given in Columns 4, 5, and 6 together with the Mach number calculated from each ratio using the well-known isentropic flow equations (given in the Handbook of Supersonic Aerodynamics[6]) with the ratio of specific heats, $\gamma=1.4$. Column 7 of Table I lists the "critical" temperature as estimated by inspection of Fig. 1. The value of the critical temperature shows a linear dependence on the stagnation pressure, as given in Fig. 2.

IV. DISCUSSION

The pressure measurements would indicate that at temperatures above that which is called "critical" for

[6] Handbook of Supersonic Aerodynamics, Navord Report 1488, 1, (April, 1950).

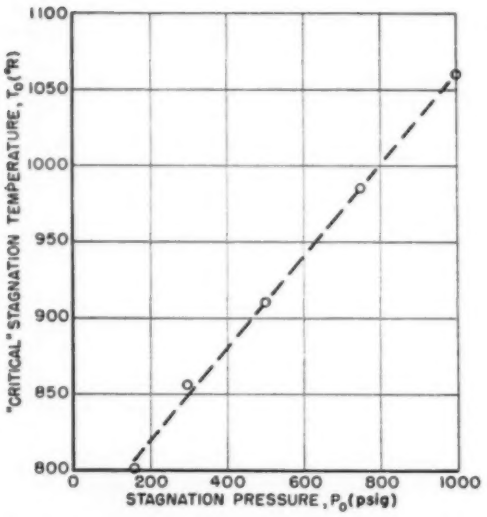


FIG. 2. Critical stagnation temperature as a function of stagnation pressure, estimated from Fig. 1.

each stagnation pressure, the flow is free from condensation effects and that furthermore, as will be discussed later, there apparently exists a state of supercooling. This phenomenon has often been anticipated; but as far as we are aware, this is the first experimental evidence of its existence in pure nitrogen that has been published.

The presence of condensation in a nozzle is typified by an apparent loss of stagnation pressure and consequently by a wide disparity between the Mach numbers obtained by the usual isentropic relations connecting the measured values of p_h, p_s, and p_0. Above the critical temperature there is, in the present experimental results, good agreement between the Mach numbers calculated by the three pressure ratios p_h/p_s, p_h/p_0, p_s/p_0. This fact leads one to believe that there is no condensation upstream of the point of measurement or at least that its effects are too small to be observed. Using the isentropic flow equations and the average of the Mach numbers shown in Columns 4, 5, and 6 of Table I, the static temperature corresponding to the critical stagnation temperature, i.e., the vapor temperature at the point of measurement, has been calculated and is shown in Column 8 of Table I. The vapor pressure as a function of vapor temperature (Columns 3 and 8) are plotted in Fig. 3. Figure 3 also shows the experimental values obtained by Cath[7] and Henning and Otto[8] for the vapor pressure of nitrogen in the range of pressures from the triple point to 1 atmosphere. For this range of pressures the solid line is calculated from the following best fit to the experimental points (given in the International Critical Tables[9]).

$$\log_{10} p_{mm} = (-344.6376/T°\text{K}) + 7.577 - 0.00476T°\text{K}.$$

The experimental values for the solid phase are given by Aoyama and Kanda;[10] the dashed line is derived theoretically by the same authors. A comparison of the results of the present tests with those obtained under equilibrium conditions shows that the very rapid expansion of the nitrogen has in each case produced about 30 Fahrenheit degrees of supercooling. The extent of the supercooling decreases slightly (~5°F in the range of 52°–64°R) with increasing vapor pressure. A state of supercooling is reported by Stever and Rathbun[4] for vapor pressures above the triple point. Rough calculations made by us on their data, as presented in their Fig. 28, have given us the point indicated as × on Fig. 3. This point is thought to represent approximately the critical vapor pressure and temperature for their work; if so, it is in good agreement with the trend shown in the tests reported here. It should be noted that reference 4 is concerned with the condensation of the nitrogen component in dry air rather than that of pure nitrogen.

No attempt has been made as yet to examine quantitatively the pressure measurements in the regime of stagnation temperatures less than "critical." A few general trends may be noted however. It is seen from Fig. 1 that the spread of the data increases with increasing mass flow; this is made somewhat more significant by the fact that the error in measuring the pressures decreases with increased mass flow. This increase in spread may correspond to an increase in instability of the flow as the density increases and may also be a factor in the small reduction of the degree of supercooling at higher vapor pressures. It may also be

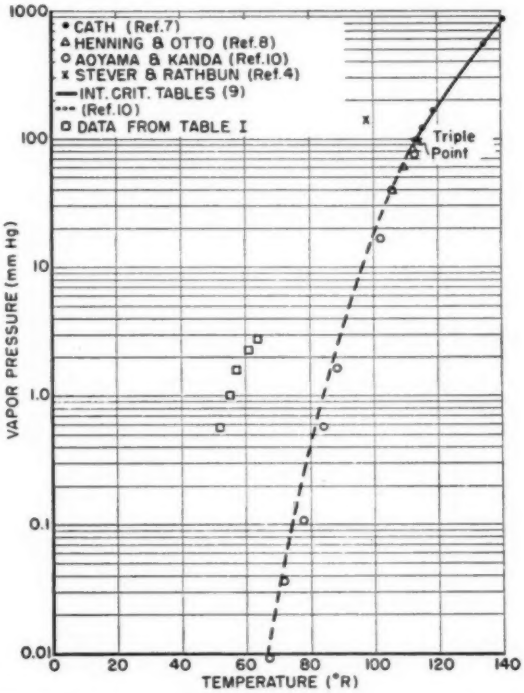


FIG. 3. Equilibrium vapor pressure of nitrogen, and measured vapor pressures.

[7] P. G. Cath, Proc. Roy. Acad. Sci. (Amsterdam) **21**, (1919).
[8] Von F. Henning and J. Otto, Physik. Z. **37**, 633–8 (1936).
[9] ICT III, 203 (1928).
[10] S. Aoyama and E. Kanda, Sci. Reports, Tohoku Imperial University Sendai. **24**, 107–115 (1935).

noted in the cases where there is the most data (e.g., the 300 psig runs) that there is a rise in pitot pressure just before the critical temperature is reached. For the cases of lower stagnation temperatures the condensation will be fully developed nearer the throat and thus in the time ($\sim 2\times 10^{-4}$ sec) required for the flow to reach the test section, considerable re-evaporation of the solid or liquid particles of nitrogen may occur with consequent stabilization of the flow before it reaches the probe. When the onset of condensation is in the vicinity of the probe, no such stabilization can occur and the effect is likely to be seen in the pressure measurements. The static pressures measured on the wall do not show as much scatter as those measured in the center of the stream and this is to be expected due to the presence of boundary layer.

A slight systematic trend may be noted in the Mach numbers of the "stable" region (Columns 4, 5, and 6 of Table I) as the stagnation pressure is increased. This is probably due to both the change in the boundary layer thickness (which is known to decrease for increasing Reynolds number) and to the variation of the gas characteristics (ratio of specific heats, gas constant, velocity of sound, etc.) under the conditions of temperature and pressure with which we are dealing. Both these factors are being studied.

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 1 JANUARY, 1952

Investigation of Stress-Strain Relations of Metal Wires by Electrical Resistance Changes

IRWIN VIGNESS
Mechanics Division, Naval Research Laboratory, Washington, D. C.
(Received July 20, 1950)

A method is described by which the stress-strain curve of a wire tensile specimen can be derived from measurements of its changes of electrical resistance as a function of strain. The method can be applied only to a selected group of metals that have experimentally been shown to be suitable. A material is suitable when its resistance changes per unit of elastic and per unit of plastic strain are each constant and different from each other. The method assumes uniform strain throughout the length of the specimen, which limits conditions studied to those prior to necking. Examples are shown illustrating the linear resistance-strain relations and the applications of the method for several suitable materials. The resistance-strain characteristics of several materials that are not suitable are given, together with a statement of some mechanisms responsible.

INTRODUCTION

IT is the purpose of this study to show how stress-strain relations of metallic wire specimens subjected to tensile stresses can be derived from measurements of the changes of the electrical resistance of the specimens as a function of strain. The legitimacy of the derivation rests upon an empirical background, and the method can be applied only to a selected group of materials. The principles involved for this derivation are as follows:

a. Linear stress-strain and resistance-strain relations are assumed to exist for elastic deformation of the specimen.

b. A linear resistance-strain relation which is of different slope than that of the elastic case is required to exist for plastic strain.

c. If the above relations are known, and if a curve is obtained which shows the change of resistance of a specimen as a function of total strain, then the elastic and plastic components of strain can be calculated, and the required values of stress can be determined as the product of the modulus of elasticity and the elastic component of strain.

d. It has been tacitly assumed that the strain and stress do not vary along the length of the specimen. This limits the amount of strain to conditions which do not involve necking. Other factors, such as strain rates, temperature, and time of load application, will require experimental study to determine their effects upon the above assumptions. It will initially be necessary to make simultaneous stress-strain-resistance measurements to determine their practical validity for the particular material and experimental conditions.

The advantage of this method is that small specimens can be subjected to strain, at some known and simple function of time, and such characteristics as the relative amount of elastic and plastic strains, the values of yield points, and other factors contained in a stress-strain curve, can be obtained from a simple oscillographic record of the resistance of the specimen.

A partial review of the background of knowledge relating the electrical resistance changes of metallic wire specimens to strain is necessary to show what general states and conditions of a metal are required for it to be a suitable specimen. This review divides itself naturally into two parts: conditions that are characteristic of elastic, and conditions that are characteristic of plastic strains. For elastic strains most metals have resistance-strain curves with maximum slopes* intermediate between $+6$ and -11. Except for some anomalous materials, as nickel,[1] the elastic region of the resistance-strain curves is linear. At the present time

* The resistance value is plotted nondimensionally as the change of resistance per unit of resistance. The slope of this curve, which is the per unit change of resistance per unit of strain, is often called the "gage factor," as it represents the sensitivity factor for the bonded-wire strain gages.

[1] E. W. Kammer, and T. E. Pardue, Proc. Soc. Exptl. Stress Analysis, **7**, No. 1, 7 (1949).

the resistance changes caused by elastic strains provide little information as to mechanisms of interest to research workers in the solid-state field. However, there has been a considerable correlation between the resistance changes and the mechanisms involved in plastic strain, so that it has often been possible, by measuring the change of resistance, to determine what particular mechanism is involved, or when a change of state occurs.

The electrical resistance of a metal is increased by changes that disturb the uniform periodicity of its lattice structure.[2] If the periodicity were exact, ideally the resistance might be expected to be zero. Such factors as thermal motions, localized strains, grain boundaries, structural imperfections, impurities, substances in solid solution, and other causes of inhomogeneities, all lead to increased resistance.

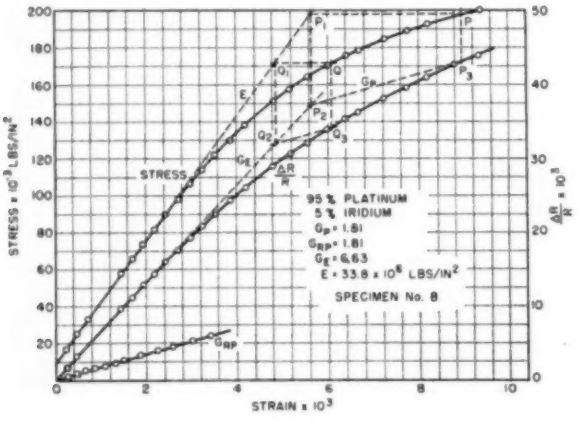


FIG. 1. Simultaneous stress- and resistance-strain curves of a platinum-iridium alloy with constructions for obtaining the plastic gage factor G_P.

The principal causes for resistance changes during plastic deformations are:

a. Multiplications of grain boundaries and crystallite imperfections.

b. Destruction of an ordered arrangement of metallic compounds or superlattices.

c. Orientation of crystallites that are nonisotropic electrically.

In addition to these effects, nucleation, precipitation, and phase-change phenomena may occur either simultaneously with, or be accelerated by, plastic flow to cause changes of electrical resistance.

It has been shown that plastic working of pure metals, constructed of crystallites with nondirectional electrical properties or of metals with impurities or alloying elements randomly disposed, causes little change in their specific resistance compared with their initial resistance at room temperature.[2a,3a] Generally a maximum change of about 2 percent occurs when a wire of such a material is drawn through a die. For the much less severe working involved in tensile tests, no definite change of specific resistance is usually noted. Thus it may be concluded that a multiplication of grain boundaries and dislocation type imperfections at most changes the specific resistance by a few percent. This would also include any possible effect of the internal elastic strains stored by this work. It would be probable, therefore, that these types of materials—which include iron, aluminum, copper, silver, gold, platinum, and many of their alloys—would be suitable for studies of this kind.

The resistance changes caused by destroying the ordered arrangement of atoms of a metallic compound may be in the order of several hundred percent.[4] Plastic working of an ordered alloy may therefore be expected to show changes in specific resistance up to this magnitude.

Andrade and Chalmers[3] performed some most interesting experiments—on cadmium and tin, in particular, whose crystals have different electrical resistance properties along their different axes—in which they were able to show that, for small plastic strains of wire specimens, there was no rotation of the crystal crystallites. The rotation usually began, however, after several percent strain. They also were able to determine whether the mechanism involved in the rotation was a shear slippage or a twinning and to show that for cadmium at high temperatures shear predominated while at very low temperatures twinning was the major mechanism.

It is probable that only those materials exhibiting very small changes of specific resistance during plastic work will be suitable specimens for stress-strain studies from resistance-strain measurements. This excludes materials that possess an ordered arrangement and those that are electrically nonisotropic.

DETERMINATION OF ELECTRICAL ELASTIC AND PLASTIC CONSTANTS OF STRAIN (ELASTIC AND PLASTIC GAGE FACTORS)

On the basis of the previous discussion it can be reasonably assumed that metals of the cubic crystalline system, which are either pure or are alloys in a disordered state, would be suitable for stress-strain determinations by means of resistance-strain measurements. Other materials that have been severely cold worked without subsequent anneal would also generally be suitable. These types of materials would have very small changes of specific resistance over the plastic range.† At the present time nickel is the only metal known not to have a linear relation between resistance and strain in the elastic region. This anomaly is probably related to its magnetostrictive properties.[5]

[2] F. Steitz, *Physics of Metals* (McGraw-Hill Book Company, Inc., New York, 1943), first edition, chapter 21. [a] p. 320.

[3] E. N. da C. Andrade, and B. Chalmers, Proc. Roy. Soc. (London) **138**, 348 (1932). [a] p. 373.

[4] Charles S. Barrett, *Structure of Metals* (McGraw-Hill Book Company, Inc., New York, 1943), first edition, pp. 243–246.

† These statements refer to the plastic behavior of the materials; proper elastic behavior is also required.

[5] S. Arzibyschew and W. J. U. Juschakow, Z. Physik **64**, 405-410 (1930); and S. Arzybaschew and V. Jushakow, Z. Physik **86**, 521–522 (1933).

Simultaneous measurements of stress, strain, and resistance were made for wire specimens as described by Kammer and Pardue.[1] The loads were applied slowly, and conditions corresponded to the usual conception of "static" tests.

Figure 1 represents a typical stress-strain and resistance-strain curve for a material having different linear elastic and plastic resistance-strain relations. The nomenclature given on these curves is defined as:

E = elastic or Young's modulus.
$\Delta R/R$ = resistance change of the specimen per unit of resistance.
$G = \Delta R/R/\text{strain}$.
$G_E = (\Delta R/R/\text{strain})_{\text{elastic}}$ = elastic gage factor.
$G_P = (\Delta R/R/\text{strain})_{\text{plastic}}$ = plastic gage factor obtained without releasing the load.
G_{PR} = residual plastic gage factor defined as above, except that it was obtained by stretching the wire beyond its yield point and releasing the load so that the resistance and strain were measured under conditions of no load.

The change of resistance of a specimen is related to the change of specific resistance by the equation

$$G = \frac{dR}{R} \Big/ \epsilon = (1+2\nu) + \frac{\partial \rho}{\rho} \Big/ \epsilon,$$

where ϵ is the strain, ν is Poisson ratio, and ρ is the specific resistance. For plastic conditions, assuming both no change in volume (i e , $\nu = 0\,5$) and zero change of specific resistance, the gage factor, G, is 2. For elastic conditions, for $\nu = 0.3$, and for zero change of specific resistance, $G = 1.6$.

For the platinum-iridium alloy of Fig. 1‡ there is therefore a relatively large increase of specific resistance for elastic strains and a very small decrease for plastic strains, provided the above assumptions are true. The points on curve G_{RP} were obtained by removing the load at various intervals and measuring the remaining strain and resistance change. The plastic gage factor, G_P, was determined as follows. For any value of stress, as PP_1, the elastic component of strain was determined from the intersection of this value of stress and the extension, E, of the elastic section of the stress-strain curve. The ordinate of P_2 is then the resistance change caused by the elastic component of strain, and that of P_3 is the total resistance change. The slope of the line, P_2P_3, is the plastic gage factor, G_P. For this example the slope, G_P, of any line such as P_2P_3, or Q_2Q_3 obtained in a similar manner, and the slope G_{RP}, all had identical values. The equality of G_{RP} and G_P is reassuring, though not a necessary condition, for the determination of the stress-strain curve by this method. As the method is empirical, it is best to determine G_P in the manner in which it is used, which would require the $PP_1P_2P_3$ construction.

‡ On all figures the strain is taken as zero when all the removable weights have been subtracted from the load; the remaining load contributes to cause the stress at zero strain.

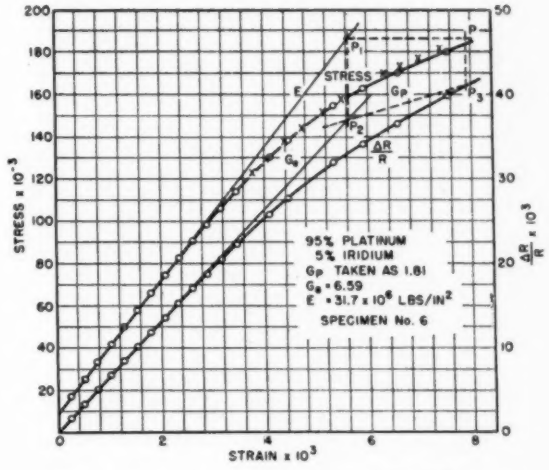


FIG. 2. Comparison of a measured and a derived stress-strain curve.

CONSTRUCTION OF THE STRESS-STRAIN CURVE

From simultaneous measurements of stress and resistance as functions of strain, as shown in Fig. 1, the components of resistance changes caused by the elastic and plastic components of strain can be determined, the slopes being G_E and G_P, respectively. Conversely, if G_E, G_P, and E are known, the stress curve is derivable from the resistance curve. This construction is shown in Fig. 2. It should be noted that, while the relative values given in these curves are accurate to better than one percent, the absolute values of stress depend upon the cross-sectional area of a wire which is about 0.001 inch in diameter and which is measured to an accuracy of only about 10 percent. Later work will be concerned with larger wires and more accurate measurements. In Fig. 2 the circles and the connecting solid lines represent experimental values of stress or resistance. The crosses represent points for a stress curve as derived from the resistance curve.

To derive this stress curve the following procedure was followed. Pick any value of total strain having a location on the resistance curve such as P_3. Draw a line

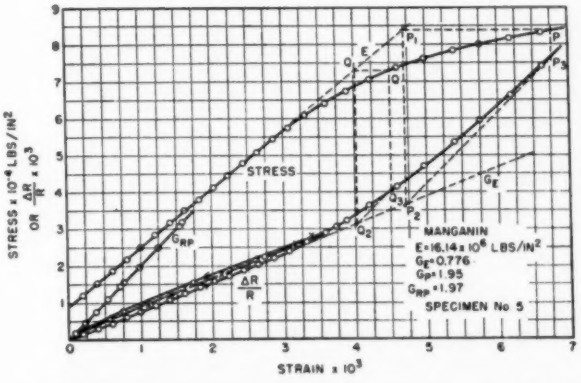


FIG. 3. Resistance-stress-strain relations for manganin. The constancy of the plastic gage factor is illustrated by the equality of the slope, G_P, of the lines Q_2Q_3 and P_2P_3.

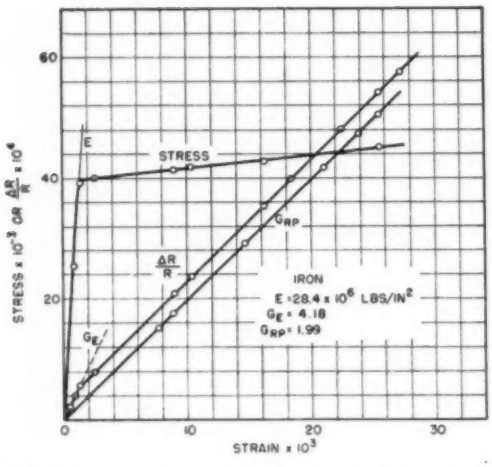


FIG. 4. Resistance-stress-strain relations for iron in which the plastic deformation is large compared with the elastic.

through P_3 having a slope G_P (which had previously been determined for this type of specimen) to intersect G_E at P_2. This gives the elastic component of strain. The corresponding stress is shown as P_1. The point P is then a point on the derived stress-strain curve.

The platinum-iridium specimens represent a condition of elastic gage factor much higher than the plastic. Figure 3 presents a condition where the elastic gage factor is smaller than the plastic. The figure also shows the construction of one point of the G_{RP} curve; the corresponding loop of the stress-strain curve has been omitted. As before the slope, G_P, of the dotted lines P_2P_3 and Q_2Q_3 are the same, which is a necessary condition for the derivation of the stress curve from the resistance curve, and the value of G_P is equal to G_{RP} within experimental accuracy.

An example of a third material is afforded by Figs. 4 and 5 for simultaneous stress-resistance-strain curves for iron-wire specimens. If the strain scale is condensed so that E or G_E becomes very steep, then the accuracy

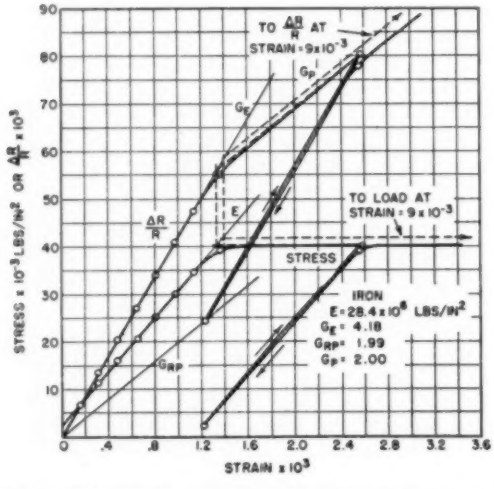


FIG. 5. Information as presented in Fig. 4 except with expanded strain and resistance scales.

of construction of the one curve from the other becomes poor. However experimental accuracy can be obtained to allow the expansion of the strain scale, as shown in Fig. 5, so as to obtain a reasonably detailed accuracy. The two derived values of G_P obtained from this figure are equal to 2.00 as compared with 1.99 for G_{RP}. Figure 5 also shows the similarity between the stress and the resistance curves as the load is gradually released and reapplied. Only a few of the experimental points are shown.

If the stress-strain curve is drawn so that the stress is zero at zero strain, then the analytical expressions of

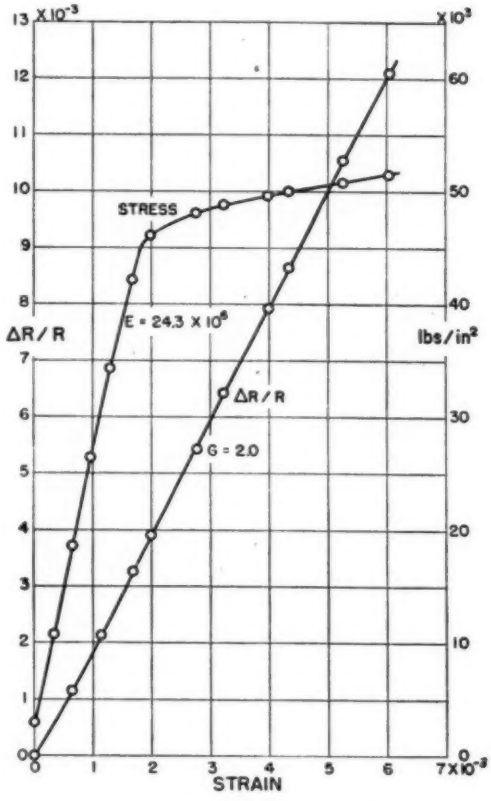


FIG. 6. Simultaneous resistance-strain relations of soft constantan illustrating the constant slope (gage factor = 2.0) of the resistance curve.

the graphical methods are as follows:

$$\sigma = E\epsilon_E$$
$$\Delta R/R = G_E\epsilon_E + G_P\epsilon_P,$$

where the subscripts E and P refer to elastic and plastic conditions, respectively, and

$$\epsilon_E + \epsilon_P = \epsilon.$$

By eliminating the strain components the stress can be expressed in terms of the three slopes and the measured values of resistance change and total strain, or,

$$\sigma = \frac{E}{G_E - G_P}\left(\frac{\Delta R}{R} - G_P\epsilon\right).$$

If the stress curve does not originate at zero strain, as is the case for the illustrated figures, then it would be necessary to add a stress equal to the ordinate intercept to the right-hand side of the last equation.

UNSUITABLE MATERIALS

There is a linear relation between the change of electrical resistance and the elastic components of strain for most metals. By coincidence this relation may be nearly the same for both elastic and plastic components. This is a fortunate coincidence when the metal is used as a bonded-wire strain-gage element, since it causes uniform sensitivity regardless of the amplitude of strain; but when this situation exists it is not possible to derive a stress-strain curve from the resistance-strain relation. Figure 6 illustrates for Constantan (soft) this condition of constant gage factor, i.e., slope of resistance-strain curve, throughout the elastic and plastic range. For best accuracy for stress-strain derivations, it is required that there be a considerable difference between the elastic and plastic gage factors.

Nickel is a unique material with respect to its electrical resistance changes as caused by elastic strains. As shown in Fig. 7, its gage factor is not constant over the elastic region of strain. The gage factor, or slope of the resistance-strain curve, is approximately -11 after the wire has been made taut by sufficient stress. The slope then gradually increases through zero and approaches a value close to 2 before the yield point of the material is reached. It has been shown (5) that this behavior is temperature sensitive and disappears at the Curie point. It has not been noticed for iron and its alloys. Such characteristics are not suitable for the stress-strain derivations.

SUMMARY

A method has been described by which the stress-strain curve of a wire tensile specimen can be derived from measurements of the change of its electrical resistance as a function of strain. The method can be applied only to a selected group of metals that have been shown by experiment to be suitable. A metal is suitable when its resistance changes per unit of elastic strain and per unit of plastic strain are each constant and are different from each other. If the strain is a known function of time, then the information for the required stress curve can be obtained oscillographically as a simple measurement of the change of electrical resistance of the specimens. The method, of course, assumes uniform strain throughout the length of the specimen. This limits experimental conditions to those excluding travelling stress waves of appreciable amplitude and to those prior to necking of the specimen. The method appears to be best suited for studying stress-strain relations over a wide range of strain-rate and temperature conditions. The higher values of strain rate, involving experiment time less than a second, would generally be most convenient from a measurement standpoint.

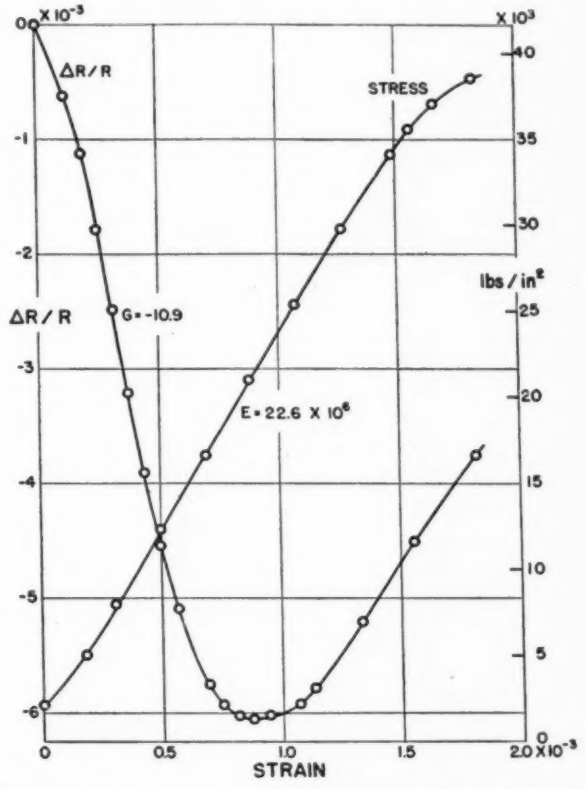


FIG. 7. Nonlinear resistance-strain relations over the elastic range of strain which is characteristic of nickel.

ACKNOWLEDGMENTS

Much appreciation is due Dr. E. W. Kammer and Mr. T. E. Pardue, who have done most of the experimental work reported here.

Glass Cracking Caused by High Explosives*

E. M. Pugh, R. v. Heine-Geldern, S. Foner, and E. C. Mutschler
Department of Physics, Carnegie Institute of Technology, Pittsburgh, Pennsylvania
(Received July 13, 1951)

High-speed photographs have been obtained of the fracture of glass produced by the detonation of a high explosive charge. Using photoelastic methods, the shock waves set up in the glass can also be photographed. In addition to providing a high intensity loading method, the explosive also serves as a convenient time base for determination of the rates of propagation of shock waves and cracking.

The results presented in this paper extend those obtained by other investigators who used low energy impacts to produce fracture.

INTRODUCTION

THE mechanical failure of glass under impact has been studied intensively by Barstow and Edgerton[1] in this country, and by Schardin and Struth[2] in Germany. The first group produced fracture by shooting spring-accelerated steel balls against glass plates; the second, by shooting rifle bullets against similar targets. Although the impacts produced by these two methods are of different orders of magnitude, and in spite of the fact that the two groups of investigators used glasses of entirely different origins and hence presumably of different chemical and mechanical properties, both of them observed a maximum velocity of crack propagation of 5040 ft/sec (about 1540 m/sec), with a probable error not exceeding two percent. This limiting velocity, according to Barstow and Edgerton, is also independent of the thermal history of the glass sample, applying equally to annealed and to highly prestressed glass.

Schardin and Struth make the further observation that minute local defects ("Kerbstellen") in the glass can originate cracks when the disturbance, traveling at the velocity of sound (5500 m/sec), passes over them. They find that such secondary fractures ("Sekundärbrüche") are initiated at a velocity of about 3000 m/sec; i.e., between that of crack propagation and sound.

It was considered desirable to extend these investigations to impacts of higher energy, such as those produced by detonating a high explosive charge in direct contact with the glass sample. Schardin and Struth[2] actually attempted to photograph cracks in glass initiated by this method, but discontinued these attempts when the accompanying experimental difficulties became apparent.

FIG. 1. Typical experimental setup showing a dummy explosive charge resting on two glass plates.

EXPERIMENTAL PROCEDURE

The investigators aforementioned used high intensity sparks to obtain photographic exposures. Because these sparks were of sufficiently short duration (1 microsecond or less) to stop all motion in the photographs, no synchronized shutter was required. High explosives, however, remain intensely luminous for periods of several hundred microseconds after detonation, so that photographs of nonluminous high speed phenomena induced by high explosives require a synchronized shutter of sufficient speed. The Kerr cell, an electro-optic device which has been described quite frequently in the recent literature,[3] can be used as such a high speed shutter.

The photographic equipment and procedure used by this group in the study of explosive phenomena has been described recently in this journal.[4] Briefly, this equipment consists of a 25KV-Kerr cell synchronized with the phenomenon and an exploding wire light source (peak intensity 4×10^8 candle power). The phenomenon itself is silhouetted against a translucent diffusing screen made of tracing paper, illuminated from behind by the

* This work was carried out under a research contract with the Office of the Chief of Ordnance, Contract No. DA-36-061-ORD-7.

[1] F. E. Barstow and H. E. Edgerton, J. Am. Ceram. Soc. **22**, 302–307 (1939).

[2] H. Schardin and W. Struth, Glastechn. Ber. **16**, 219–231 (1938).

[3] At the time of writing, the most recent article appears to be that by Quinn, McKay, and Bourque, in J. Appl. Phys. **21**, 995–1001 (1950).

[4] Pugh, Heine-Geldern, *et al.*, J. Appl. Phys. **22**, 487–493 (1951).

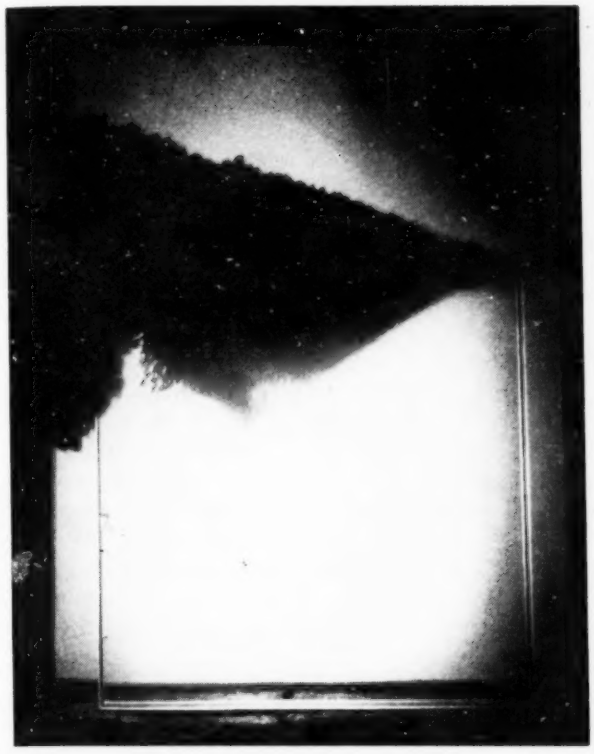


FIG. 2. Detonation of Primacord in contact with a $6\times6\times1\frac{1}{4}$-in. glass plate. Detonation proceeds from left to right at the rate of 6420 m/sec.

exploding wire. By this method, shadowgraphs are obtained at exposure times of about $\frac{1}{2}$ microsecond.

The explosives which were used for the tests to be described below were Primacord, a high velocity fuse with a detonation velocity of 6420 m/sec, and Pentolite, one of the most powerful high explosives, with a detonation velocity of 7510 m/sec.

The glass used for this work was commercial (soda-lime) plate glass obtained from the Pittsburgh Plate Glass Company. Unless otherwise stated, all test targets consisted of annealed plate glass of dimensions $6\times6\times1\frac{1}{4}$ inches.

Figure 1 shows a setup similar to that used in several of the experiments. It consists of two glass plates, each measuring $6\times6\times1\frac{1}{4}$ in. supporting a cylindrical explosive charge with an electric detonator protruding from the rear end.

DESCRIPTION OF PHOTOGRAPHS

For the photograph shown in Fig. 2, a length of primacord was laid along the upper face of a $6\times6\times1\frac{1}{4}$-in. plate glass target. The Primacord was detonated from the left, and at the instant of exposure, detonation had progressed to the right-hand vertical face of the target. The dark region above the glass plate is caused by the opaque explosive products, expanding radially at a velocity of 2250 m/sec. Although these explosive products are luminous, their luminosity cannot compete with the brilliant background. The dark region in the glass plate is caused by intense shattering damage. From the known detonation velocity u of primacord (6420 m/sec) and the angle α(33.5°) between the advancing cracking front and the upper edge of the glass plate, the velocity v of advance of the cracking front can be determined from

$$v = u \sin\alpha.$$

In this case, the velocity v is found to be 3500 m/sec.

Still referring to Fig. 2, it will be noted that hardly any individual cracks can be discerned. This is mainly caused by the fact that shattering is quite complete. It may be of interest that none of the glass plates survived this type of experiment, the largest recoverable fragments not exceeding a few cubic millimeters in volume. Another observation to be made on Fig. 2 is the fact that the angle α, measured as 33.5°, is constant over a limited distance only. Close to the upper right-hand corner of the plate, the angle is considerably smaller, indicating that the initial velocity of propagation of the cracking front is very much smaller.

In later paragraphs, evidence will be presented to show that the high cracking rate of 3500 m/sec is caused by the continuous initiation of cracks on the part of the tension wave produced by reflection of the

FIG. 3. Similar to Fig. 2, but the explosive used here is Pentolite with a detonation velocity of 7510 m/sec. In the original photograph, a shock wave is faintly visible near the upper right-hand corner of the glass plate.

FIG. 4. The outline of the shock wave, which is faintly visible in Fig. 3, is brought out more clearly by means of photoelastic methods.

shock wave from the two 6×6-in. faces of the glass plate. Close to the explosive charge, no reflected tension wave has yet been formed and the observed cracking velocity is much smaller.

A shock of still higher intensity can be produced if the Primacord is replaced by Pentolite. Figure 3 shows a cylinder of Pentolite (diameter 1.6 cm) detonating in contact with a glass plate similar to the one of Fig. 2. Detonation again proceeds from left to right, and at the instant of exposure has progressed slightly beyond the right-hand vertical face of the target. Figure 3 differs qualitatively from Fig. 2 in two respects: first, the detonation products of Pentolite compete in luminosity with the very bright background; second, the shock imparted to the glass plate is sufficiently intense to produce a shock wave that is faintly visible without the aid of photoelastic methods. This shock wave, making an angle of 55° with the upper target surface, propagates at a velocity of 6150 m/sec, thus exceeding the velocity of sound in glass by at least 10 percent. Even more remarkable is the fact that, close to the source of shock, the cracking front is approximately parallel to the shock wave and hence propagates at about the same velocity as the shock, although lagging slightly behind it.

The shock wave can be brought out more clearly by resorting to photoelastic methods. For that purpose, a sheet of polaroid was placed behind the glass plate shown in Fig. 4, and the combination of target and polaroid was rotated until the polaroid axis was crossed with that of the first Kerr cell polaroid. In other respects, the setup in Fig. 4 is identical to the one in the preceding figure.

The shape of the shock wave in Fig. 4 is exactly as predicted by a Huygen's construction, while the uniformly bright area immediately behind it is not so easily explained. It appears that there exists a region of fairly uniform strain between the shock wave and the cracking front. The shock velocity deduced from this picture is again 6150 m/sec, in agreement with the result obtained from Fig. 3.

The cracking patterns shown in Figs. 3 and 4 indicate that close to the source of shock, where the shock intensity is highest, cracks are originated with the velocity of the shock wave. As the shock intensity decreases, the outline of the cracking front becomes more irregular and falls behind the advancing shock wave.

Figure 5 shows the same experimental arrangement as that shown in Fig. 2 (Primacord in contact with glass), except that photoelastic methods were again employed in order to render the shock wave visible. For the weaker Primacord, the observed shock velocity is about 5800 m/sec, as compared to 6150 m/sec for Pentolite. However, in the case of this weaker shock, the

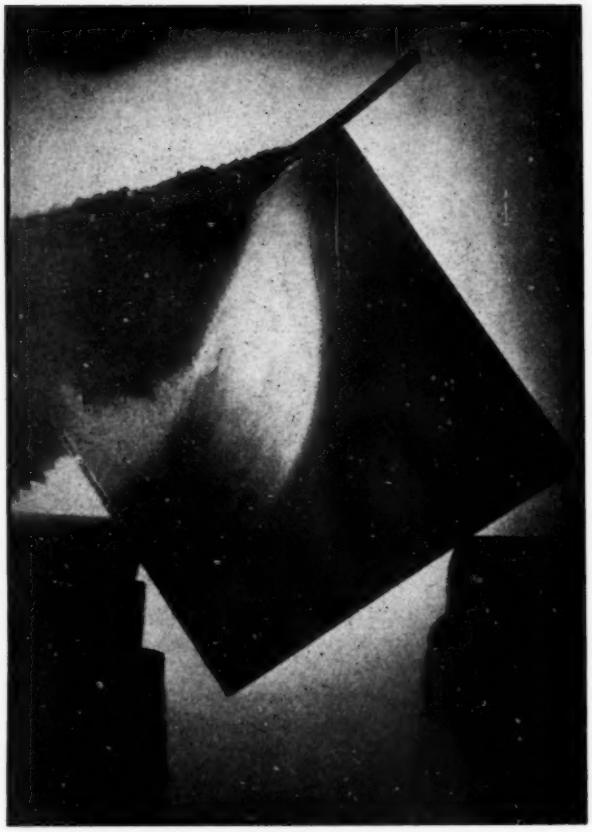


FIG. 5. Photoelastic methods applied to the setup shown in Fig. 2 (Primacord placed against glass) again bring out the shock wave caused by the detonating explosive.

cracking front no longer advances at the same rate as the disturbance; in this case, the velocity at which fracture is initiated is about 3400 m/sec, which agrees reasonably well with the value of 3500 m/sec deduced from Fig. 2.

At this point it may be profitable to consider the manner in which a shock wave will behave in a finite block of glass. The shock waves shown in Figs. 3 to 5 are obviously compression waves, since they were imparted to the glass by the detonation of an adjacent explosive charge. Upon reaching a free boundary, such as a glass-to-air interface, a tension wave will be

FIG. 6. Instantaneous cracking pattern produced in a glass plate by a Pentolite cylinder standing on the upper surface. The glass plate is viewed through its narrow (6×1¼ in.) surface.

reflected back into the glass block. The question naturally arises whether the cracking damage shown in Figs. 2 to 5 was caused (directly or indirectly) by the original compression wave or by the reflected tension wave originating at the front and back surfaces of these targets. This question is answered by the result shown in Fig. 6, which was obtained by standing a cylindrical pentolite charge on a glass plate which is viewed through its narrow (6×1¼ in.) surface. The contour of the shattered region is such that it could only have resulted from the tension waves reflected from the two sides after passage of the original compression wave. Unfortunately, the disturbances producing

FIG. 7. Stack of two glass plates with a cylinder of Pentolite laid along the uppermost 6×6-in. surface.

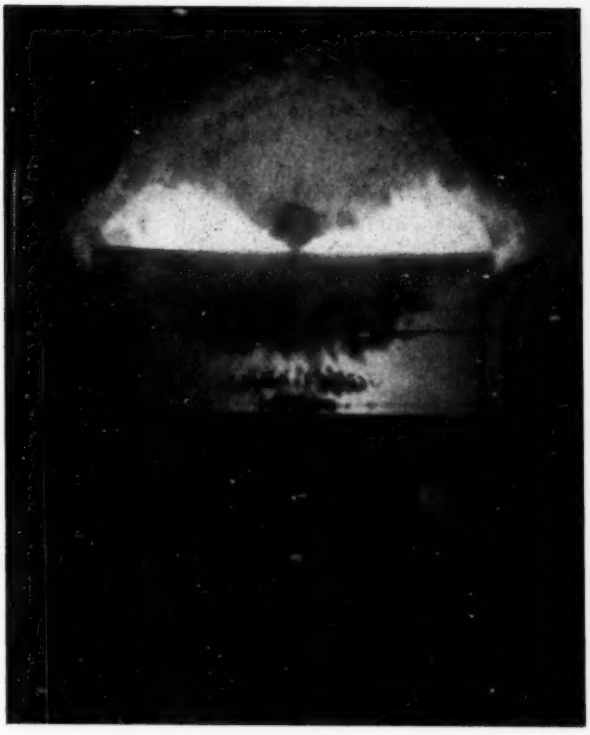


FIG. 8. The setup of Fig. 7 rotated through 90 degrees, so that detonation of the Pentolite proceeds toward the observer.

FIG. 9. Primacord detonated along the edge of thin ($\frac{1}{8}$ in.) glass plate having parallel scribe marks on one of its surfaces.

cracking damage could not be made visible by photoelastic means, because the narrow faces of the available glass plates were of comparatively poor optical quality.

In an attempt to delay formation of reflected tension waves, two glass plates similar to the ones shown above were stacked as shown in Fig. 7, with their $6\times1\frac{1}{4}$-in. surfaces toward the camera. A cylindrical Pentolite charge (diameter 20 mm) was laid along the center of the uppermost 6×6-in. surface. In the photograph, detonation of the charge has progressed nearly to the right-hand edge of the stack. The original shock wave, clearly visible in this picture, is seen to be partially transmitted from the first to the second plate without appreciable delay, and partially reflected back into the first plate at the interface.

Probably the most significant observation to be made from Fig. 7 is the relatively low propagation rate of the cracking front near the explosive, amounting to about 2000 m/sec, as compared to 6150 m/sec deduced from Figs. 3 and 4. It must be concluded that the higher rate results from the early formation of reflected tension waves in the case of the experiments illustrated by these figures, and that in the absence of such tension waves, as in Fig. 7, the cracking front proceeds with a lower velocity.

This conclusion also helps to explain an observation made earlier on Fig. 2; cracking there proceeds at the rate of 3500 m/sec except very close to the explosive charge, where, in the absence of tension waves reflected from the broad surface the velocity is considerably lower.

Three other cracking fronts can be observed in Fig. 7; each of these can be correlated at least qualitatively with the disturbance (pressure or tension wave) producing it. In general, it appears that the weaker the disturbance (i.e., the farther it has traveled from the source of shock), the slower the corresponding cracking front will propagate and at a given point the longer the delay between the passage of the disturbance and the first appearance of visible cracks.

Figure 8 illustrates the same experiment as Fig. 7, except that the setup has been rotated 90°, so that detonation of the charge now proceeds toward the observer. This picture shows clearly the discontinuous nature of the cracking visible in the bottom plate. The reflected tension wave seems to initiate cracks more or less simultaneously in a number of different locations.

This type of discontinuous fracture, first observed by Schardin and Struth,[2] can be induced artificially by intentionally weakening the glass with surface scratches, as illustrated in Figs. 9 and 10. The experiment illustrated in these pictures consisted of the detonation of Primacord laid along the edge of a thin ($\frac{1}{8}$ in.) glass

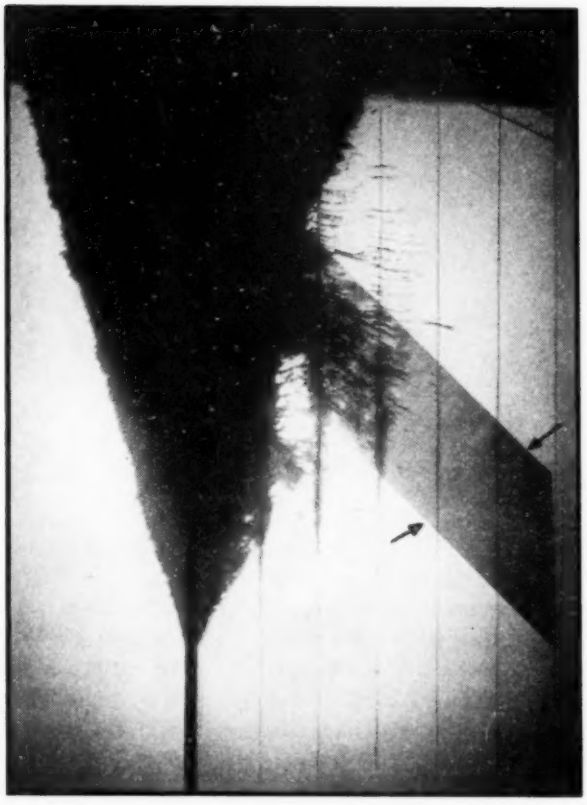


FIG. 10. Photoelastic methods applied to the experiment in Fig. 9. The arrows, on either side of the diagonal strip of polaroid, point out the shock wave which is only faintly visible in the original photograph.

plate after parallel scribe marks had been made on one of its surfaces. Both pictures show how cracks are initiated discontinuously and far ahead of the regular cracking front in those places where the glass had been weakened by surface scratches. Part of the shock wave was made visible in Fig. 10 by a strip of polaroid. If more of this shock wave were visible, one would observe that it initiates cracks at the scribe marks without appreciable delay.

SUMMARY

The results presented so far can be partly summarized as follows:

(1) The velocity of shock waves in glass increases with increasing shock intensity. The highest observed shock velocity is 6150 m/sec, which exceeds the velocity of sound by at least 10 percent.

(2) "Secondary fracture" of glass, first observed by Schardin and Struth, is produced very readily by detonation of a high explosive in contact with the sample. For sufficiently intense disturbances, this secondary fracture is initiated at the velocity of propagation of the disturbance. The time delay between passage of the disturbance and the first appearance of cracks decreases with increasing shock intensity.

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 1 JANUARY, 1952

The M.I.T. Linear Electron Accelerator*

P. T. Demos, A. F. Kip, and J. C. Slater
Laboratory for Nuclear Science and Engineering and Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts

(Received August 2, 1951)

The M.I.T. linear electron accelerator consists of a circular wave guide fitted with uniformly spaced irises, and sustains a standing-wave, transverse magnetic π-mode oscillation at a resonant frequency of 2800 Mc/sec. The phase velocity is equal to the velocity of light. Electrons are injected into the accelerating tube at an energy of 2 Mev by a standard Van de Graaff generator. Acceleration up to an energy of at least 18 Mev has been obtained, using the 21-foot length of the accelerator. Peak accelerated currents of the order of 10 ma, and average currents of the order of 1 μa, are obtained. The accelerator is fed by 21 tunable magnetrons (Raytheon RK5586), phased in synchronism. In this paper, the principles of design, constructional features, and operation are discussed. The accelerator is now being used for nuclear experimentation.

I. INTRODUCTION

THE M.I.T. linear electron accelerator† was started soon after the war, in the Research Laboratory of Electronics at M.I.T., with the aim of applying microwave methods to the problem of accelerating electrons. The object from the beginning was to determine if it was possible to build an accelerator of great length, capable of producing very high energy electrons; but the accelerator actually built was designed to be as short as possible consistent with testing the principles. The result has been the development and construction of an accelerator about twenty-one feet long, powered by twenty-one pulsed magnetrons (Raytheon RK5586) operating at about 10.7-cm wavelength, of a type which could be extended to indefinite length. This accelerator takes electrons at 2 Mev, injected from a Van de Graaff generator, and accelerates them to approximately 18 Mev, producing a well-defined beam of approximately a microampere average current (the order of 10-ma peak current), with a homogeneity in energy of the order of magnitude of ten percent. Thus, the feasibility of this particular design has been demonstrated; and at the same time the accelerator which we have built proves to be a useful instrument for nuclear studies. To make use of it, the accelerator was transferred late in 1949 to the jurisdiction of the Laboratory for Nuclear Science and Engineering at M.I.T., where it is now being used for photodisintegrations, photofissions, and other types of nuclear experimentation.

This paper, with several others which are being published at about the same time,‡ contains the major information about the design of the accelerator, the interesting technical steps that led to various features of its development, and the study of the characteristics of its beam. Many persons have been associated at one time or another with its development, design, and construction; it has been throughout a cooperative undertaking. It did not seem feasible to list all of these contributors as joint authors of the present paper; therefore, we list here the persons who have made major contributions to it: W. H. Bostick, R. J. Debs, P. T. Demos, E. Everhart, I. Halpern, J. Halpern, A. F. Kip, M. Labbitt, L. C. Maier, Jr., S. J. Mason, I. Polk, R. A. Rapuano, J. C. Slater, and J. R. Terrall. One of the present authors (J. C. S.) has been in general charge of

* This work has been supported in part by the ONR, AEC, Signal Corps, and Air Materiel Command.

† See bibliography at end of this paper, particularly reference 11.

‡ See references 5, 6, and 14.

the project; two others (P. T. D. and I. H.) are now supervising the operation of the accelerator.

The thinking resulting in the accelerator design proceeded in the following way. To make an indefinitely long accelerator obviously requires an indefinite series of power sources, spaced along the length of the accelerator, and operating in synchronism. The simplest way to synchronize many power sources is to use power amplifiers, fed from a common signal. At the time the design was begun, no amplifiers of sufficiently high power were available in the microwave region; and since our aim was to study the acceleration of electrons and the use of existing microwave equipment, rather than to develop new microwave power sources, we used the existing high power generators, the magnetrons. It is, of course, well known that the accelerator project at Stanford in the mean-time has developed high power klystron amplifiers. If they had been available at the time, we should presumably have used them, since they seem to provide the most straightforward method of feeding an accelerator. Nevertheless, our alternative solution by means of magnetrons is practicable, and may have some useful applications in future accelerator design.

A magnetron is a self-excited oscillator, whose frequency is usually determined by the resonant circuit built into it. It is clear that to operate many magnetrons in synchronism, their frequencies must be sufficiently close together so that no two will get out of phase with each other by an appreciable angle during the pulse, which is two microseconds, or approximately 6000 cycles of the oscillation at the 2800-Mc/sec frequency used in our case. Thus, the frequencies must agree, say, to a few parts in 100,000. Clearly, this cannot be achieved without tunable magnetrons; fortunately, the Ratheon magnetron mentioned above is tunable. The necessary accuracy can hardly be achieved, however, even with tunable magnetrons, without some form of frequency stabilization, since ordinarily the frequency of a magnetron drifts by more than this amount because of changing temperature, changing power level and load, and so on.

The best way to stabilize the frequency of a magnetron is to couple it to a very high Q resonant cavity. This coupling cannot be accomplished without having a resistive load which absorbs part of the power; otherwise, the oscillation proves to build up a very high level in the magnetron itself, without putting appreciable power into the resonant cavity. Part of the power of the magnetron is used in the resistive load, part goes into the stabilizing cavity. It is difficult to adjust the parameters so that more than about half the output power goes into the resonant cavity. As a result of this situation, we have a choice of two possible methods of applying this method to feeding power into the accelerator.

The essential feature of the accelerator is a wave guide, capable of sustaining a wave with longitudinal electric field, traveling with the same velocity as the electrons, so that the electrons can always be in the same phase of the electric field, continuing to gain energy if they are in proper phase. This demands a traveling wave. We can either make the wave guide nonresonant, by using a resistive termination, so that it will allow a traveling wave in only the direction in which the electrons travel; or we can make it resonant, by using reflecting terminations, so that we set up a standing wave consisting of traveling waves in the two opposite directions. In our design, we have chosen to use the standing wave or resonant design, which allows us to use the accelerator cavity itself as the stabilizing cavity for the magnetrons. Our object, then, is to feed as much power as possible into this cavity; and we find that we can feed about half the power of the magnetrons to the cavity, the remaining half being dissipated in the resistive load.

It is worth pointing out that the alternative solution would have been possible, using a traveling-wave nonresonant accelerator, using it as the resistive load for each magnetron, and furnishing each magnetron with a separate frequency-stabilizing resonant cavity, properly tunable so that they could all be tuned together with the requisite precision. Our solution appealed to us as rather the simpler, in that it avoided the construction of these stabilizing cavities. It is also worthwhile pointing out that for an accelerator fed by power amplifiers, in which frequency stabilization is not necessary, the advantages are all in favor of the traveling-wave nonresonant accelerator, since it furnishes a frequency-insensitive load which can be fed easily from an amplifier and since it does not dissipate part of the power in the traveling wave proceeding in the direction opposite to the electrons, which, of course, is useless for acceleration.

Once we had decided to make the accelerator resonant, it was necessary to study the spectrum of resonant modes. Since we were injecting electrons of 2-Mev energy, which are therefore traveling very close to the velocity of light, we wished to build an accelerator in which the phase velocity of the wave was just the velocity of light; investigation of the electron dynamics showed that this would bunch and accelerate a considerable fraction, about $\frac{1}{4}$, of all injected electrons. The simplest accelerator for this purpose proved to be a circular wave guide fitted with transverse irises with circular holes; it is easy to choose the dimensions of such a loaded guide so as to slow down the phase velocity to the velocity of light. If a finite length of such a guide is terminated at both ends by reflecting walls midway between irises (of course pierced by holes near the center to allow the electron beam to enter and leave), it has a set of resonant modes; if there are N identical sections, there will be $N+1$ modes. In the mode of lowest frequency, the electric field will travel in the same longitudinal direction in each section between successive irises, corresponding to an infinitely long wavelength;

in the mode of highest frequency, there will be a reversal of phase from one section to the next, so that this is called the π-mode, on account of the phase difference of π from one section to the next. The spectrum of the $N+1$ modes shows that some modes are bunched close to the frequency of the mode of infinite wavelength, others close to the π-mode, with a wider mode spacing in the middle. The frequency separation between the mode of infinite wavelength and the π-mode depends only upon the properties of a single section, becoming smaller as the holes in the irises become smaller. On the other hand, the number $N+1$ of modes increases with the length of the accelerator, so that the average frequency spacing between adjacent modes is proportional to $1/N$; and it can be shown that near the edges of the band, where the modes bunch together, the frequency spacing between adjacent modes is proportional to $1/N^2$. Thus, it is clear that for a very long accelerator the modes can come indefinitely close together.

This clearly imposes a limit on the use of the accelerator cavity for frequency stabilization, for a magnetron or other self-excited oscillator can stabilize on a mode of a resonant cavity only if no other modes are too close in frequency. It is thus impossible to make an indefinitely long accelerating cavity and use it to stabilize magnetrons. Since our object was to find how to design an indefinitely long accelerator, it is obvious that this had to be made up of a number of isolated sections, each short enough to have adequate frequency separation between modes for magnetron stabilization, operating independently from an electrical point of view, but with holes big enough for electrons to travel from one section to the next, though small enough to present the passage of radiofrequency power. Our final design consists of seven such isolated sections, each approximately three feet long, separated by "drift tubes" holes to admit the electrons from one section to the next. It is clear that to make an accelerator out of such electrically independent sections, we must have means of tuning each of the sections to the same frequency, to the order of the few parts in 100,000 mentioned earlier, so that all parts of the accelerator will operate at exactly the same frequency; and we must have means of operating the sections in phase with each other, so that the electrons will meet an rf field of the appropriate phase in each section.

As we have stated before, the main purpose of the present development was to encounter, in a short accelerator, the type of problem which would have to be overcome in a very long accelerator. For this reason we operated in the π-mode, in which the mode separation is a minimum. Study of the mode separation, and the requirements for stabilization, led us to think that a section 24 half-wavelengths long would have an adequate frequency separation between the π-mode and the next adjacent mode, whereas any longer section would be marginal. When such sections were set up, however, unexpected mode difficulties appeared. The type of mode in which we operate has a longitudinal electric field, or is a TM mode. However, we found that the lowest type of TE mode of the cavity, which of course also has a spectrum of $N+1$ resonant modes, lay in the same general frequency range as the desired TM modes, and by particularly bad luck, for the case $N=24$, one of the modes of this TE series lay almost in exact coincidence with the π mode of the TM series. We investigated the whole spectrum of TE modes, to see how they were located for different values of N, and found that for certain values of N the nearest TE modes were at a safe distance from the TM π-mode, whereas for other N's they were dangerously close. The largest value of N for which the separation was safely large was $N=16$; and for that reason we standardized on that value.

For practical purposes we wished to construct the accelerator out of sections about a foot long; it was convenient to space the magnetrons this closely without having them get in one another's way. Each of the one-foot sections is fitted with a wave-guide input for a magnetron, together with the necessary resistive load to ensure its operation properly into the resonant load. The "one-foot" dimension is, in fact, six half-wavelengths, so that for the 16 half-waves of the basic unit of the accelerator, we need two of these one-foot cavities, and a third cavity consisting of four half-waves and a drift tube one wavelength long. One-foot cavities of these two varieties, standard and drift-tube, were accordingly made, each containing a magnetron input, and the accelerator was assembled from two standard cavities, a drift-tube cavity, and so on, repeated through its length. Each three-foot unit, operating by itself electrically, then had to be tuned, and this was accomplished by built-in tuners. The three magnetrons operating into one such stabilizing cavity automatically lock to the cavity and to each other, so that they operate in complete synchronism; we are then left with the problem of making all seven three-foot units phase together.

The first requirement for this phasing is simply to tune all seven cavities to the same frequency, with the required accuracy of a few parts in 100,000; this proved not to be difficult with the tuners. The frequency equality can be tested by a beat method to be described presently. The resonant frequency of the cavities was found to drift slowly as the temperature changed; and at first the dissipation of power in the cavities resulted in a progressive temperature increase, which brought about a corresponding frequency drift, with the necessity of frequent tuning. This was obviated by water-cooling the cavities, and with this water cooling the frequency equality can be maintained for hours at a time, with very little adjustment.

The second requirement for phasing is that all cavities start their oscillation in the same phase. It is known that if a relatively weak signal is fed to a self-excited oscillator while it is building up its oscillations, the

oscillation will build up with a definite phase relation to the incoming signal. This is in the first place the way in which the three magnetrons fed into a single resonant unit get into synchronism: one of them is pulsed slightly before the other two and provides the signal which establishes the phase of the other two. A similar method is used to phase the seven three-foot units with respect to each other. One of the sections is pulsed slightly before the other six, and serves as a master section. From this section, by coaxial lines from its resistive load, small samples of its output power are fed to each of the other six sections, building up a large enough power level in each such section before its oscillation starts to make it operate in the proper phase. A TR box is provided at the place where each coaxial line is fed into one of the sections, to break down and protect the coaxial line from overloading when the magnetrons in such a section fire. A line stretcher is provided in each coaxial line, so that the time delay between the master section and each other section can be adjusted, resulting in the correct relative phase for each section.

The adjustment of each section is conveniently accomplished by having a probe output from each section by means of which a sample of its power is fed to the control bench. This sample can be observed on a spectrum analyzer, and the tuners adjusted by remote control to get approximately the desired frequency. The sample can also be observed on an oscilloscope displaying the envelope of the power in the cavity, as it builds up during the two microsecond pulse, dying down again in approximately the same time. By combining the outputs of two three-foot units and observing on this scope, one can easily check all aspects of the phasing and tuning. If the two units are correctly phased together, the combined pattern of the two is the same for each pulse, and if the lines from the probes to the scope are properly adjusted in length, the envelopes of the two units will add, whereas if the phasing is absent the phase will sometimes be right for the signals to add, sometimes for them to subtract, so that the pattern will show characteristic jittering behavior. If the two units are not only phased, but tuned together, the envelopes will add throughout the approximately four microseconds of excitation, while if they have different frequencies, there will be beats between then, which can be at once observed. This furnishes the method mentioned earlier for checking the frequency adjustment.

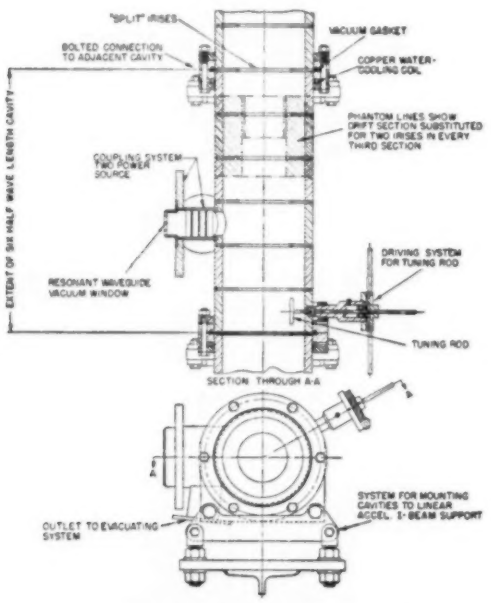

"SPLIT" IRISES
VACUUM GASKET
BOLTED CONNECTION TO ADJACENT CAVITY
COPPER WATER-COOLING COIL
PHANTOM LINES SHOW DRIFT SECTION SUBSTITUTED FOR TWO IRISES IN EVERY THIRD SECTION
COUPLING SYSTEM INC. POWER SOURCE
EXTENT OF SIX HALF WAVE LENGTH CAVITY
DRIVING SYSTEM FOR TUNING ROD
RESONANT WAVEGUIDE VACUUM WINDOW
TUNING ROD
SECTION THROUGH A-A
SYSTEM FOR MOUNTING CAVITIES TO LINEAR ACCEL. I-BEAM SUPPORT
OUTLET TO EVACUATING SYSTEM

FIG. 1. Scale drawing of the linear accelerator tube.

It is clear that there is nothing in this construction or method of testing which limits the length to that which we have used. It is for this reason that the design could be adapted to an indefinitely long accelerator. There is no reason why a single master phasing signal could not be applied to the phasing of as many magnetrons and resonant sections as desired, and the test procedure outlined could be used to adjust them all to synchronism. It is to be noticed that a magnetron, when controlled by a phasing signal in this way, has many of the properties of a power amplifier, so that this solution of the accelerator problem is not so different from the power amplifier solution as might be thought at first sight.

The injected electron beam, as has been stated, is delivered at an energy of 2 Mev from a Van de Graaff generator. It is a pulsed beam, giving approximately a 1-microsecond pulse injected at the time when the cavity amplitude is at its maximum. The practical detail of getting the pulsing signal, which, of course, is essentially the same signal used to trigger the magnetrons, up to the high voltage terminal of the generator has been solved by letting the pulsing signal flash a strobotron lamp, which shines through a window in the wall of the pressure chamber of the generator, illuminating a photocell in the high voltage terminal, and hence actuates a circuit which momentarily debiases the electron gun in the high voltage terminal. Currents of some 40 milliamperes are at present being delivered by the gun, of which about $\frac{1}{4}$, or 10 milliamperes peak, is accelerated to the end of the accelerator; with the duty cycle of approximately 1 in 10,000, this gives an average current of about a microampere. The gun at present being used is limited to this current by space charge, but nothing in principle forbids the use of a considerably higher beam current.

The energy of the accelerated beam is analyzed by a magnetic deflection method. For adjustment of the accelerator, it is found useful to have an alternating magnetic deflection, which deflects the beam alternately to left and right; the highest voltage will then be achieved when the separation of the two deflected images is a minimum. These deflected images, observed by letting the electrons impringe on a fluorescent screen, are seen visually by telescope and mirrors from the

control room. The line stretchers regulating the phases of the various three-foot units are then adjusted to give maximum electron energy. It is found that with practice this adjustment can be made without great difficulty.

The entire length of the accelerator is shielded magnetically by high permeability metal, to protect the beam from the magnetic field of the earth and the magnetron magnets. The only focusing which is required is that provided by the focusing coils in the Van de Graaff generator, which present a rather sharp spot at the input of the accelerator. Several sets of Helmholtz coils in the first few feet of the accelerator, set so as to deflect the beam transversely, aim it correctly down the axis of the accelerator and keep it on the axis. After the first few feet, it continues along the axis without trouble and emerges at the far end as a fairly sharp spot, though with a certain amount of transverse jitter and a certain amount of diffuseness surrounding the spot, some of which tends to be in a horizontal direction. We suspect that this horizontal deviation may be tied up with remnants of the *TE* field, which would be in a horizontal plane since the wave guides feeding in the power have a horizontal *E*-field direction. When a well-defined beam is needed, as for good energy measurement by the magnetic analyzer, the beam is passed through a collimating slit, but this is not needed for ordinary purposes.

We have stated that about $\frac{1}{4}$ of the incident beam is accelerated to the far end. The remaining $\frac{3}{4}$ is almost all lost in the first few feet of the accelerator, as we find by comparing the current at the far end with the current after passing through a few feet of accelerator. The electrons which are lost are those which are not in the phase to be accelerated, but are rather in the phase to be decelerated. It is recalled that there is a radical defocusing field in a linear accelerator, which is important for slow particles, but which becomes inoperative as the particles approach the velocity of light. Presumably, the particles which become decelerated are going so slowly that they become defocused and hit the sides of the accelerator before they have a chance to go far. There is no evidence of unaccelerated particles getting through to the far end of the accelerator. These relatively slow electrons striking the sides of the accelerator produce a strong x-ray emission along its length, so that it must be placed in a shielded room and operated entirely by remote control. The radiation emitted from the target end of the accelerator is particularly strong, consisting both of high voltage x-rays, and, since the electron energy is above the threshold for emission of neutrons from many common elements, of copiously emitted neutrons. For this reason the target is very heavily shielded.

The maximum energy of the accelerated electrons is about 18 Mev. This is considerably less than was hoped when the accelerator was designed; but when all the factors are considered, it proves to be in accordance with theory. The respects in which the operation is less satisfactory than the initial estimates are the following: the Q of the cavities is not quite as high as hoped; the magnetrons deliver considerably less power, both because we find that they cannot be operated at as high current as was hoped originally and because the standing wave ratio of about 2 into which they operate, on account of the resistive load and stabilizing cavities, reduces their power output more than was hoped at first; and on account of the finite length of the pulse, the voltage in the cavities does not have a chance to build up to its full steady-state value. When all these features are considered, the observed electron energy and also the breadth of the energy peak check expectations well. We believe that if better magnetrons were available, they could be used with this same accelerator and produce considerably greater acceleration.

FIG. 2. "One-foot" six half-wavelength cavity showing the frequency tuning rod.

We now go on to detailed discussion of most of the features of the accelerator construction and design.

II. THE ACCELERATING CAVITY

The basic "one-foot" cavity (six half-wavelengths long) of which the accelerator is constructed is shown in the drawing and photographs of Figs. 1 and 2; the drift tube section is indicated by the "phantom" lines in Fig. 1. The dimensions are as follows: internal diameter of cylinder, 4.000 in.; distance between irises, on centers, 2.106 in., equal to half the wavelength of 10.698 cm, for operation in the π-mode; thickness of irises, 0.125 in.; diameter of hole in irises, 2.506 in., the edges of the holes being rounded to semicircular cross section. The resonant wavelength of the properly terminated cavity, as mentioned above, is 10.698 cm.

so that the traveling associated with the π-mode travels with the velocity of light. The cavity ends with a split iris at each end, half as thick as a regular iris; since no electric current must travel from one-half of the iris to the other, it is immaterial whether the two halves of the iris make electric contact or not. Each cavity is flanged at each end to allow it to be joined to its neighbors with bolts, and is provided with an O-ring in a groove for vacuum seal. The cavities are made from oxygen-free high-conductivity copper (O.F.H.C.) which was machined in the form of separate cylinders and irises. The component parts were then joined together by silver soldering in a hydrogen atmosphere.

The simplest type of termination which will allow such a finite cavity to have the same resonant frequency

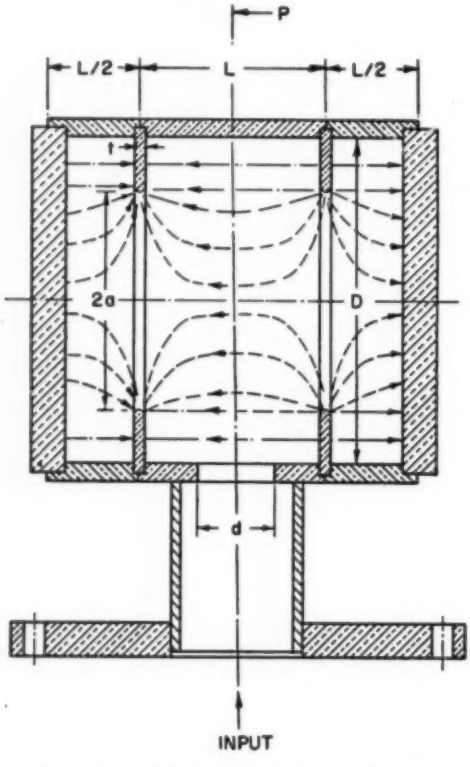


FIG. 3. Experimental one-wavelength cavity.

as an infinite length of cavity is a conducting plane closing the cross section midway between two irises. By symmetry the lines of force automatically strike such a plane normally, so that it does not disturb the field pattern. The preliminary experimental rf work to establish the exact cavity dimensions was done using the cavity of Fig. 3, terminated in this way. The resonant frequency of such a cavity was determined for a number of different values of dimensions, in the neighborhood of those computed from the theory of reference 10. The dimensions finally chosen, quoted above, agree closely with those predicted by the theory, and are the result of considerable experimentation.

Where the cavity joins to a drift tube, the cavity termination would be a plane midway between irises, as in Fig. 3, if it were not for the perturbation produced by the drift tube opening. Considerable experimentation was required to find how much the plane had to be displaced from the midplane to compensate for this perturbation. Some perturbation was also produced by the coupling holes which admit the power, and the dimensions stated compensate for this. When a three-foot unit is assembled out of these one-foot sections, it should have a resonant wavelength for the π-mode of just 10.698 cm, and a wave pattern which varies sinusoidally in going down the cavity. In reference 5 it is shown how the pattern can be tested experimentally. Such tests were made on all the sections, and showed in some cases that the excitation was too strong at one end or the other. This is the type of error to be expected from slightly incorrect position of the drift tube termination. It was found possible to correct these errors by inserting small collars inside the drift tubes, projecting very slightly into the cavity, and adjusted empirically to produce a satisfactory wave pattern. The machining tolerances were such that when a three-foot section was assembled, the actual wavelength was close enough to 10.698 cm so that the tuners could tune it to exactly that value.

The tuning method, shown in Figs. 1 and 2, consisted of a copper rod, $\frac{1}{8}$ in. in diameter by 1 in. long, arranged so as to perturb the electric field of the cavity, inserted into each one-foot section. The rod may be rotated in a plane perpendicular to the radius by means of its supporting shaft, which is brought through a vacuum seal in the cavity wall. In this position the rod perturbs the cavity frequency least when lying in a direction perpendicular to the cavity axis, or the main component of the electric field. Using one such tuner in each one-foot section, the tuning range was made sufficiently large (about 0.25 percent) without perturbing the field at the axis of the cavity excessively. The three tuners of each of the three cavity sections are ganged together and adjusted during operation of the accelerator by remote control.

Provision for feeding power to the cavity was made by using a feed hole in one of the center half-wavelength sections, power being brought to the feed hold through a short piece of 10-cm rectangular wave guide fitted to the cavity, flanged at its open end for coupling to the power source, and provided at the plane of the flange with a glass window which was electrically matched. Orientation of the wave guide input was of course such that its E-field direction would be parallel to the axis of the cavity, thereby giving preference to the excitation of TM modes. It was desired to make the feed holes of such size that they would not reflect back any of the power fed in from the magnetrons, in steady operation. The size was determined by feeding the attenuated output of a klystron oscillator to the cavity through a standard 10-cm slotted section, and measuring the standing wave ratio. For this purpose, each of the

cavities was terminated at its open ends with quarter-wavelength sections fitted with split irises, but otherwise similar to the end sections of the experimental one-wavelength section of Fig. 3. Thus when such a test was made, the cavity was seven half-waves long, whereas it was desired to have the iris matched for a section six half-waves long. From this it was easy to calculate the standing wave ratio on resonance which was desired. Beginning with a hole of diameter $\frac{7}{8}$ in., the diameter was increased in small steps until enough data were available to allow extrapolation to the correct feed-hole size to give the desired standing wave ratio. This value was then used for the feed-hole diameter in all of the cavities which followed. A measurement over a number of finished cavities at a later date, however, showed them to be mismatched; the input standing wave ratio, measured by the method above, was found to be from 3 to 5 db in magnitude. This was explained by the tarnishing of the cavity interiors which had occurred in the meantime, lowering the unloaded Q of the cavities and thereby decreasing the coupling between the power source and cavity. Assuming that the tarnishing had proceeded to its practical limit, it was decided to open the feed holes further to match to the new conditions of the cavity surface. The resulting hole size became 1.205 in., and is the value in use at the present time.

The same measurement of standing wave ratio of the terminated cavities permitted a measurement of the unloaded Q, by techniques described elsewhere (see references 3 and 12). The value of Q_0 was found to range from 16,000 to 19,000, lying between 17,000 and 18,000 for most of the cavities. As is seen later, these values of Q_0 predict an electric field intensity in the cavities, in terms of the estimated power input, which is essentially in agreement with direct measurements of the energy of the electron beam.

III. THE JUNCTION FROM MAGNETRON TO CAVITY

The feeding of a self-excited oscillator into a resonant load, which is to operate as a stabilizing cavity, is a relatively complicated problem, and we shall sketch the essential features of the method used in this section. The reader is referred to references 11 and 12, where the essential parts of the problem are taken up. In Sec. 10 of reference 11, and in Sec. 9.3 of reference 12, and in particular in Fig. 39 of that section, it is shown how, by plotting both electronic and circuit admittance on the same plot, we can determine the frequency and power level at the point of steady state operation into a resonant load, and at the same time (as shown in Sec. 9.4 of reference 12) can discuss the build-up of oscillation. The example used in Fig. 39, and in the discussion, is that actually encountered in the design we have used for the linear accelerator. The resonant circuit, consisting of cavity and resistive load, is effectively a series combination of the matched cavity impedance and resistance equal to the characteristic impedance of the line. We shall now describe this combination more in detail, and indicate how it works.

We recall that if we look through a wave guide and an iris into a resonant cavity, there is a particular plane of reference in the wave guide, called the plane of the detuned short, across which the impedance can be written approximately in the form

$$\frac{Z}{Z_0}=\frac{1/Q_{ext}}{j((\omega/\omega_0)-(\omega_0/\omega))+(1/Q_0)}, \tag{1}$$

where Z is the impedance, Z_0 the characteristic impedance of the wave guide, Q_{ext} the external Q of the cavity (the Q describing the power loss through the coupling of the cavity to the wave guide), Q_0 the unloaded Q, ω the angular frequency at which measurement is being made, and ω_0 the resonant frequency of the cavity. In our case, if we take a one-foot section, terminate it so that it is six half-wavelengths long, and measure its input impedance looking in through the wave-guide input, we have adjusted the cavity to match; that is, we have opened the iris, as described in the preceding section so that Q_{ext} for such a length of cavity equals Q_0, so that when $\omega=\omega_0$ we have $Z=Z_0$. The plane of the detuned short was determined for this cavity by the type of standing wave measurements described at the end of the preceding section and found to be very closely coincident with the plane of the iris in the wall of the cavity.

Let us now put in series with the cavity described above a matched resistive load, so that the impedance of the combination has Z/Z_0 equal to unity plus the value given above for the cavity alone. If we take the reciprocal of this combination, to get the admittance, we find that it has the form shown in Fig. 39 of reference 12, so that this combination represents the case taken up in that reference. As shown there, particularly on p. 199, this gives a stabilizing circuit which causes the oscillation to build up properly in the resonant cavity, and to have a frequency determined by ω_0, the resonant frequency of the cavity. It is this combination which we have used. The effect of the resistive load can be described simply in words. In Secs. 5.4 and 9.4 of reference 12, we show that, though the steady-state input impedance of a cavity is as given in Eq. (1), its impedance during build-up is very different, and can be determined by the device of replacing the angular frequency ω in Eq. (1) by a suitably chosen complex frequency. In particular, at the start of oscillation the impedance is zero; the cavity acts like a short circuit. If we tried to operate a magnetron directly into such a cavity, it could not start, for a self-excited oscillator cannot operate into a short circuit. The matched series resistance means that at the start of build-up the magnetron sees a matched load, and starts smoothly, at first sending most of its power into the resistive load. As the cavity excitation gradually increases, however, its impedance builds up; and by the time a steady

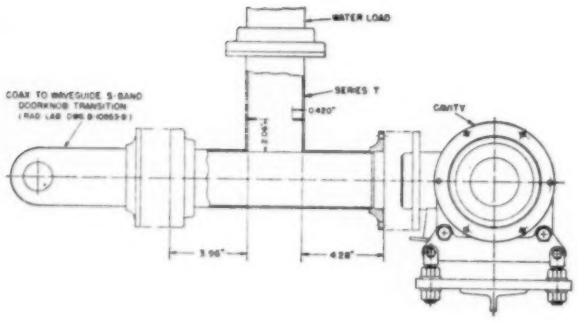


FIG. 4. Exploded side view of magnetron transition, series T, and cavity.

state is reached, it has a matched impedance. At this time the magnetron is seeing a load of $2Z_0$, consisting of the resistive load and the cavity in series, and half the power is flowing into each.

The situation which we have just described, though by no means the only way of setting up a stabilizing cavity, is a practicable one, and it is the one which we have adopted. We must next ask how it can be experimentally realized We are first directed to get a matched resistive load; this is easily done by a suitable designed water load in a length of wave guide. Next we must combine this load in series with the cavity impedance of Eq. (1), and then we must present this series combination as the load of the magnetron. The simplest way to combine two microwave loads in series is by means of a symmetrical wave guide T, and such a T is used with the linear accelerator in the manner shown in Fig. 4 to couple a water load, Z_0, in series with each of the cavities.

The theory of such a T is given in Sec. 7.1, reference 12. In Eq. (1.3) of that section we show that the relation between the ratios of impedance to characteristic impedance looking out from the three arms of the T, which we may call Z_1, Z_2, Z_3 (where Z_1 and Z_2 refer to the two arms looking straight through the T and Z_3 refers to the arm at right angles) are related by an equation

$$Z_3+\alpha(Z_1+Z_2)=-jX_0, \qquad (2)$$

where α and x_0 are two constants determined by the properties of the T and where this simple relation holds only across certain definitely determined planes of reference in each arm of the T. Directions are given in that section for finding these planes of reference, by standing wave measurements, and also for finding α and X_0. Furthermore, it is shown in that section how a T can be modified, by introducing suitable irises across its arms, so as to make $\alpha=1$ and $X_0=0$ in Eq. (2). All these steps were carried out in the design of the T's for the accelerator. Then if arm 1 represents that leading to the accelerator cavity, arm 2 that leading to the magnetron, and arm 3 that to the matched water load, the impedance seen by the magnetron, which is $-Z_2$, will be Z_1+Z_3, or the required series combination of the two.

In setting up this T, it is essential to remember that the planes of reference must be properly chosen. Thus the plane of reference in arm 1 must coincide with the plane of the detuned short of the accelerator cavity, or be a whole number of half-guide wavelengths from it. The plane of reference in arm 3, attached to the water load, is not significant, since the load is matched. The plane of reference of the T in arm 2, attached to the magnetron, must coincide with, or be a whole number of half-wavelengths from, the plane of the detuned short of the magnetron. This plane was determined by cold test carried out on the magnetron. In other words, all dimensions of the T assembly are critical, and relatively small deviations of any of them can completely prevent proper operation. The dimensions of the T satisfying the electrical and physical requirements are given in Fig. 4, and apply, of course, only when used in conjunction with the appropriately dimensioned wave-guide couplings to the magnetron and the linear accelerator cavity shown in the same figure. To arrive at the final T dimensions, cold tests were run on a T designed strictly on the basis of the measurements described above. The experimental T was terminated on the appropriate arms with a water load and cavity, and the standing wave ratio of the power fed into the T was measured for a range of frequencies about resonance. The T dimensions were then "trimmed" in the way indicated by the measurements to obtain the proper standing wave ratio on resonance together with a symmetrical variation of the standing wave ratio on either side of the resonant frequency.

IV. THE PHASING TECHNIQUE

The photograph of Fig. 5 shows the essentials of the phasing system used. The phasing signal is obtained by tapping off some of the power normally dissipated in one water load of the master section. Probes are inserted into the wave guide of this water load; and rf power, estimated to be about 30 db down from the cavity power level, is fed into each of the slave sections through the system of rf cables 1 to 6. After passing through a motor controlled line stretcher, the phasing signal enters each slave section through one arm (A) of a second wave guide T introduced into the system fed by the slave section's center magnetron. To prevent useful power developed in the slave section from leaving by the same path, a TR box is used in each of the arms A which breaks down for the high field of the slave section, but not for the much lower field of the phasing signal. The TR box behaves like a short circuit when discharging and is placed in the arm A so as to be in series with the associated water load; in this way the loading of each of the center magnetrons is the same as for its section companions once phasing has taken place and the TR boxes have broken down. For the TR box to be used in this way, it is necessary to pulse the magnetrons of the master section slightly in advance (0.2 to 0.3 microsecond) of the slave sections. As

we stated in Sec. 1, to avoid competition between the three magnetrons of a three-foot section for phasing control, it has also been found necessary to prepulse one of the magnetrons of each section before the other two by about 0.2 microsecond.

The successful operation of the phasing is checked, as indicated in Sec. 1, by taking signals from the probes of both master and slave section, mixing them together, rectifying them by means of a crystal, and observing the resultant on a high speed synchroscope. The resultant amplitude of the mixed signals will clearly depend upon the relative phase of the master and slave section oscillations: if the phase relationship varies from pulse to pulse of the oscillations, the synchroscope pattern will vary in a correspondingly random way; if the sections are phased, the pattern amplitude will remain more or less constant depending upon the degree of phase locking. As may easily be shown, the variation observed in the resultant signal is a function of the relative phases of the mixed signals; that is, if there is good, but not complete, locking, the "jitter" observed in the resultant signal appears deceptively large when the mixed signals are out of phase, and deceptively small when these are in phase. For this reason, since the line lengths from the cavity probes to the synchroscope bear no particular relationship among their lengths, it has been found convenient to use an adjustable line stretcher in the line of one of the signals being mixed. Experience has shown that the locking conditions can be observed at a glance using the line stretcher; if for some reason there has been an appreciable reduction in the degree of locking, the resulting change in the synchroscope pattern is quite noticeable. The phasing, in those cases when it was found not normal, has always been completely random; the system is apparently simple enough so that failures, when they occasionally occur, are such as to cause complete loss of the phase locking. The visual method of checking the phasing has therefore proved to be quite reliable and is always used during routine operation of the linear accelerator.

While locking does occur with such relatively small phasing signals, it is natural to expect some random phase variation. In reference 12, Sec. 9.5, where the phasing of a self-excited oscillator by an external signal is discussed, it is shown that the phase angle between operation of the oscillator and the phasing signal depends upon the electronic admittance of the oscillator, which, of course, is subject to random variation due to tube noise. Thus, there are phase "jitters" which cannot be directly controlled by the phasing signal. Several direct measurements of the phase jitter in the present linear accelerator system have been made and yield a value of approximately $\pm 10°$ for the probable random phase variation of any slave section with respect to the master section. This is in good agreement with a value deduced from a sequence of measurements on the energy of an electron beam accelerated by two sections.

V. THE MODULATOR AND PULSING CIRCUIT

The linear accelerator magnetrons are pulsed by twenty-one thyratron modulator circuits (consisting of thyratron, two-microsecond pulse-forming network, and pulse transformer), each isolated by means of a radio frequency choke from a common high voltage power supply. The high voltage supply is a conventional resonant charging system designed for a repetition firing rate of 120 cycles per second, and develops a voltage up to 15 kv. Triggering of the twenty-one thyratron modulators is synchronized to occur at 120 cycles per second and at a time when the power supply charging cycle has reached its peak. The trigger pulse to the thyratrons originates in a trigger circuit actuated by a rectified 60-cycle signal tapped from the same power line as is used for the high voltage power supply, thereby automatically providing the necessary synchronization. From the trigger circuit the signal is passed to the synchroscope used for viewing the linear accelerator rf fields, going through this to a master modulator capable of delivering the necessary power for the simultaneous triggering all twenty-one of the magnetron modulators. The output of the master modulator is applied first to a pulse transformer whose primary is in series with the strobotron used to trigger the Van de Graaff generator electron gun, as described later. The pulse developed by the pulse transformer secondary is fed to a parallel line from points along whose length the signal is tapped off to the successive grids of the twenty-one magnetron modulator thyratrons. A simple adjustable RC network is provided in the grid circuit of each of the thyratrons and allows the firing time of each magnetron to be adjusted over a range of about 0.5 microsecond, which is sufficient to allow for proper phase locking of the various sections, as discussed earlier.

VI. THE ELECTRON BEAM

Although other methods of injection have been considered, only the 2-Mev Van de Graaff generator referred to earlier has been used to inject electrons into

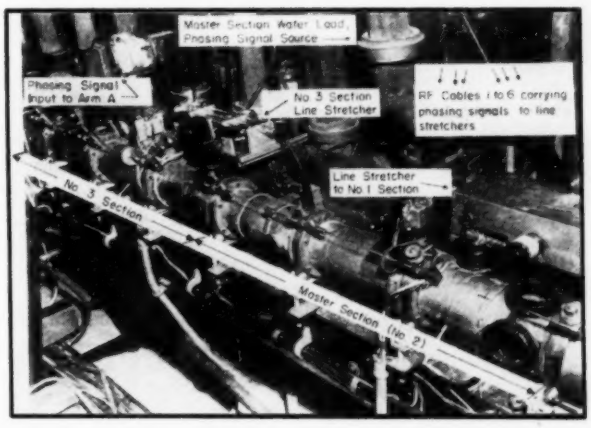

Master Section Water Load
Phasing Signal Source
Phasing Signal Input to Arm A
No 3 Section Line Stretcher
RF Cables 1 to 6 carrying phasing signals to line stretchers
Line Stretcher to No 1 Section
No 3 Section
Master Section (No 2)

FIG. 5. View of the linear accelerator showing two 3-cavity sections and phasing accessories.

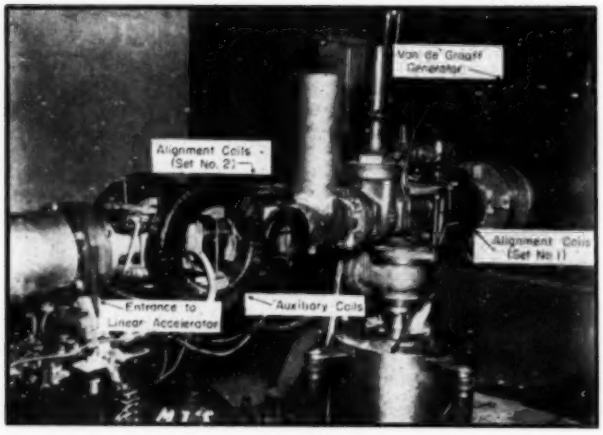

Van de Graaff Generator
Alignment Coils (Set No. 2)
Alignment Coils (Set No. 1)
Entrance to Linear Accelerator
Auxiliary Coils

FIG. 6. Coupling tube and beam alignment coils between Van de Graaff generator and linear accelerator.

the linear accelerator thus far. The generator, a standard pressurized unit,§ was adapted for the required pulsed operation by means of circuitry developed in the linear accelerator laboratory. Some mechanical and electrical difficulties which were at first experienced with the generator were eliminated by redesign of the high voltage column supports and certain features of the x-ray tube. In its improved form the Van de Graaff generator has functioned consistently and well, its only serious limitation at the present time being with regard to available beam current.

Because energy is delivered by the magnetrons to the accelerator cavities in the form of pulses it is, of course, necessary to pulse the electron source as well. We have stated that the cavities build up their excitation during the two-microsecond magnetron pulse, and then require approximately another two microseconds to die down again. For one microsecond at the center of this interval the cavity field is fairly constant, and electrons are injected during this microsecond. The variation of the intensity of the cavity field over this time introduces a certain spread in the energy of the accelerator beam, but this is not as great as the spread occurring naturally from the fact that different accelerated electrons are at different phases of the rf field.

As mentioned earlier, triggering of the current pulse is done by a photoelectric system because of the insulation problem. The thermionic filament source is normally kept at a positive voltage with respect to the focusing cup (as little as 50 volts is sufficient in the present electron gun), under which conditions no current is injected into the Van de Graaff accelerating tube. When the linear accelerator master modulator applies a pulse for triggering the magnetron modulators, a strobotron placed in series with the master modulator output transformer is energized and illuminates a photocell contained in the electrostatic dome of the Van de Graaff generator. The photocell in turn triggers a circuit consisting of an amplifier and blocking oscillator whose output, a negative voltage pulse of one-microsecond duration, is applied in series with the positive filament bias voltage and allows a microsecond burst of electrons to be admitted into the accelerating tube.

Adjustment of the electron injection time relative to the build-up of oscillations in the cavities has been made by means of a delaying condenser in parallel with the strobotron tube. The proper injection time was first determined by examining, on a synchronized oscilloscope, the voltage pulse applied to the Van de Graaff filament and the rectified output of a probe inserted in one of the cavities. The delaying condenser was then adjusted until these bore the proper relationship to each other. Later checks were made by observing the time of arrival of the pulse of current received at the linear accelerator target with respect to the cavity field amplitude and showed the previous adjustment to be correct.

The shape of the injected and accelerated beam has been observed using a fluorescent screen, both at the exit of the Van de Graaff generator and at the linear accelerator target more than twenty feet away. Using a longitudinal magnetic focusing field at the exist port of the Van de Graaff generator accelerating tube, it is possible to focus the beam to a spot of diameter $\frac{1}{16}$ in. to $\frac{3}{32}$ in. at the entrance of the linear accelerator. The same beam at the linear accelerator target, when focused by means of the longitudinal field with the rf accelerating field turned off, forms a good image of the filament which, as should be expected, has a diameter of about $\frac{5}{8}$ in.

Even when aligned by optical methods so that the exit port axis of the accelerating tube and the filament coincided with the linear accelerator axis the beam emerging from the Van de Graaff generator was found to be as much as $\frac{1}{4}$ in. to $\frac{1}{2}$ in. off axis. Two sets of vertical and horizontal deflecting coils (Fig. 6) have been used to bring the beam on axis. One set has been placed immediately after the exit port of the Van de Graaff generator and is used to deflect the beam to a point on the axis at the entrance of the linear accelerator, approximately four feet away. The second set is placed as closely as possible to the linear accelerator entrance and is used to align the beam in a direction parallel to its axis. In practice it has been found easiest to adjust the alignment coils while observing the beam after its acceleration by the cavities, choosing the setting which gives the best accelerated beam shape.

An injection peak current of 30 to 50 ma peak has been obtained using a thoriated tungsten filament identical in dimensions to the standard tungsten filament provided with the Van de Graaff generator. This corresponds, at the duty cycle of the linear accelerator (a one-microsecond pulse every 1/120 second), to an average injection current of approximately 3.5 to 6 microam-

§ Constructed by the High Voltage Engineering Corporation of Cambridge, Massachusetts.

peres. (The high energy beam, as we have stated before, after acceleration by the linear accelerator cavities, is only about $\frac{1}{4}$ of the injected beam magnitude, or approximately 1 to 1.5 μa in average value.) This current is considerably less than the current carrying capacity of the linear accelerator, but represents, because of space charge limitations, the maximum current available from the Van de Graaff generator in its present form. Modifications to produce larger beam current have been put off for the present in favor of doing some nuclear experiments with the beam at its present level.

We have just described the beam as delivered by the Van de Graaff generator, when the accelerator is not operating. The first step in turning on and adjusting the accelerator is to be sure that this unaccelerated beam can be observed on the fluorescent screen at the far end of the accelerator. The operation of turning on the accelerator and bringing it into adjustment can then be carried out in the following way. All the magnetrons are turned on, and each magnetron is tuned to the correct operating frequency region. Each of the three-foot sections is tuned to the correct frequency, and it is verified that the magnetrons of each three-foot section are locked to the cavity. The relative phasing of the three-foot sections is checked by use of the synchroscope, as described earlier. Then all that remains to be done is to adjust the relative phases of the various three-foot sections with respect to the master section, by adjusting the line stretchers so as to produce maximum acceleration of the beam. The procedure which has been found most practicable for this operation is first to adjust the relative phase of the first and second sections for maximum acceleration, using the deflecting magnet mentioned earlier to estimate the amount of acceleration. After this, the remaining sections, 3 through 7, are adjusted in succession. By the time the beam energy has become high enough, it is producing neutrons and gamma-rays at the target, and the maximization of the neutron flux proves to be the most sensitive test for making the final adjustments of phasing to obtain maximum beam energy. Finally, the magnetic fields of the several beam alignment coils along the length of the accelerator and the alignment coils between the Van de Graaff generator and the linear accelerator are adjusted to give the best beam shape. Because very little adjustment to either the rf or the beam alignment systems is needed between successive accelerator runs, the various adjustments can be gone through quite quickly. The time required for bringing the accelerator into operation, including warm-up of the magnetrons, tuning of the rf system, and beam adjustment, is usually about thirty minutes after turn-on.

VII. CHARACTERISTICS OF THE ACCELERATED BEAM

The accelerated beam forms a much narrower spot than the unaccelerated beam, when observed at the target end of the accelerator; for the unaccelerated beam, after forming a narrow spot at the entrance of the accelerator, spreads geometrically, whereas the accelerated electrons are pulled into line along the axis by the accelerating field. The main portion of the spot is about $\frac{1}{16}$ in. in diameter, with a hazy portion surrounding it, although, as we have mentioned earlier, there is a certain amount of transverse jitter, so that the illuminated area actually is of the order of $\frac{1}{4}$ in. in diameter. As we have mentioned earlier, a collimating slit is used when precise energy measurements are made by magnetic deflection.

Measurements of beam current were made by catching the beam in a collector, connected with circuits which allowed observations of both the instantaneous current pulse and the average beam current. It is by this method that the values of current quoted earlier were obtained. Measurements were made of the unaccelerated current, of the current accelerated by only one three-foot section, when the other sections were turned off, and of the current accelerated by the whole accelerator. The current accelerated by a single three-foot section was found to be about 0.38 of the unaccelerated current, and that accelerated by the whole accelerator was about 0.24 of the unaccelerated current, justifying our earlier statement that most of the electrons lost were lost in the first few feet of the accelerator.

The energy distribution of the electron beam was obtained by deflecting it with a magnetic analyzer whose field could be varied, and observing the magnitude of that part of the beam passing through a narrow collimating slit at a fixed angle to the undeflected beam direction as a function of magnetic field. The magnetic field was varied periodically and presented as a horizontal sweep on an oscilloscope, and a signal proportional to the beam deflected through the slit was presented vertically on the oscilloscope, so that a spectrum

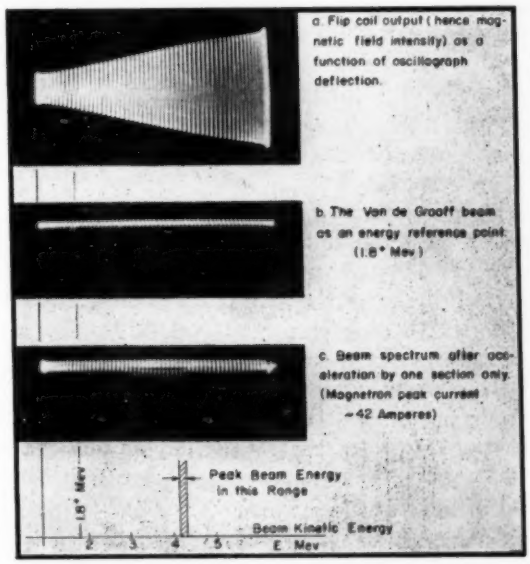

a. Flip coil output (hence magnetic field intensity) as a function of oscillograph deflection.
b. The Van de Graaff beam as an energy reference point (1.8* Mev)
c. Beam spectrum after acceleration by one section only. (Magnetron peak current ~42 Amperes)
1.8* Mev
Peak Beam Energy in this Range
Beam Kinetic Energy
2
3
4
5
E Mev

FIG. 7. A typical calibration set (a, b), and spectrum photograph (c).

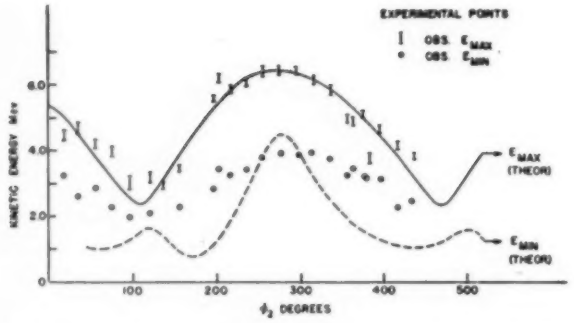


FIG. 8. Effect of variation of phase of one section on the beam accelerated by two 3-cavity sections.

of energies was observed on the oscilloscope, as shown in Fig. 7. Calibration of the energy scale was accomplished by means of a voltage generating flip coil, located in the deflecting field, whose output was applied to the oscilloscope vertical deflecting plates at intervals during each experiment. To make a tie-in with a point of known energy, the beam of the Van de Graaff generator alone was deflected through the analyzing slit and the flip coil output corresponding to this point noted. The data given in Figs. 8–10 are from photographs of such oscillograph traces and show the variation in maximum and minimum beam energy with variations in the phase of one of the accelerator sections. Because of the finite width of the collimating slit, the true maximum energy of the electron beam must lie somewhere between the maximum and half maximum points of the observed spectrum. The maximum energy accordingly appears in Figs. 8, 9, and 10 as a vertical line extending between these limits, which is a sufficiently accurate representation in view of the other errors involved in such a photographic method.

Observations were made, using the spectrum viewing technique, on the energy of the electron beam accelerated by each of the accelerating three-foot sections, and by groups of two, three, and seven sections. These observations were compared with theory in two ways. In the first place, in reference 11, p. 487, we find that the maximum accelerating field along the axis of the accelerator should be $\alpha(377PQ_0/\lambda_0)^{\frac{1}{2}}$, where α is a geometrical constant, P is the power input in watts per meter, Q_0 is the unloaded Q of the accelerator, and λ_0 is the wavelength in meters. On p. 491 of reference 11 it is stated that α for the M.I.T. accelerator, computed by the methods of reference 10, is 0.48. In reference 6 this value of α is checked by two independent experimental methods, and the computed value is verified. If then we put in for Q_0 the value of about 17,500 found for most of the experimental cavities, we find the field to be 3.76 $(P)^{\frac{1}{2}}$ megavolts per meter, where P is now expressed in megawatts per meter of cavity length. One feature of the problem, however, was neglected in this derivation: the fact that in the two-microsecond pulse the excitation in the cavity has not had quite time to attain its full steady state value. In reference 12, Sec. 5.4, the build-up of power in a resonant cavity is discussed. It is shown that under certain circumstances (not very different from those encountered in this problem) the voltage in the cavity builds up according to the function $1-\exp(-\omega_0/2Q_L)t$, where Q_L is the loaded Q, which in our case is half the unloaded Q. The quantity $\omega_0 t/Q_0$ equals the number of cycles, divided by $Q_0 2\pi$; and with Q_0 equal to 17,500, this is approximately the number of cycles divided by 2800. Since the frequency is close to 2800 megacycles per second, this means that the quantity $(\omega_0 t)/Q_0$ is almost exactly unity for a time of one microsecond, so that after two microseconds the voltage has built up to a fraction $1-\exp(-2)$, or about 0.87 of its final value. Thus, we should multiply our earlier figure by this factor, obtaining a final value of 3.28 $(P)^{\frac{1}{2}}$ for the maximum field along the axis.

From this formula, we can estimate the magnetron input power which must have been used when for instance the electrons are accelerated from 2 Mev to 18 Mev in the length of the accelerator. The active part of the accelerator (excluding drift tubes) consists of 56 wavelengths, or almost exactly 6 meters, so that the accelerating field must have been 16/6 Mev/meter $=2.67$ Mev/meter. With this value, we then find that $P=0.66$ megawatt/meter, or about 4.0 megawatts in the 6 meters, or about 190 kw supplied by each of the 21 magnetrons. This would indicate a power output of about 380 kw per magnetron, since half the power is dissipated in the water load. This probably is about what the magnetrons are actually delivering. With the two-microsecond pulse it has been found that a peak power of about 750 kw into a matched load is about the best that can be obtained from an average magnetron, and this is reduced considerably further by the operation into the mismatched load $2Z_0$ of the cavity and water load in series, both through the direct effect on the Rieke diagram, and by causing tube instability to occur at a lower anode current than with a matched load.

The second method of comparing observations with theory was by use of the method of reference 14, in which the effect of a single three-foot section on an injected electron of arbitrary energy and phase is cal-

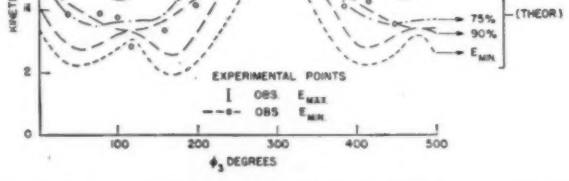


FIG. 9. Effect of variation of phase of final section on the beam accelerated by three 3-cavity sections.

culated, and shown in Fig. 2 of reference 14. That figure is calculated for an accelerating field of 2.4 Mev/meter, rather than 2.67 as found above, so that its predictions were modified to take account of this somewhat larger observed value. The course of the analysis was guided by the observations on the beam accelerated by a single section. Both the spectrum range (~3 to ~4 Mev) and the ratio of the emerging to injected beam currents are consistent with what would be predicted by the theory if all but the 35 percent to 40 percent of the electrons injected at phases 0° to −135° (on Fig. 2 of reference 14) are lost by defocusing in passing through the section. Assuming that only these electrons are available for acceleration by further three-foot sections, we can use the methods of reference 14 again to find the distribution of energy after passing through different numbers of sections. In Figs. 8–10 we show theoretical values of maximum and minimum energy of the beam after acceleration by additional sections of the accelerator, referring to energies attained by members of the group with injection phase range 0° to −135°. Figures 9 and 10 show also the ranges of energy below maximum occupied by (a) 75 percent and (b) 90 percent of this same group and afford a qualitative, but rapid, estimate of the energy distribution.

The experimental and adjusted theoretical data show quite good agreement except for the values measured for the minimum spectrum energy. These last disagree with the predicted values at all but the optimum phase settings of the two- and three-section experiments. The disagreement is not surprising, however, since the theoretical variations assume that all those electrons emerging from the first section are accelerated without loss from that point on. The experience with the first section itself, in which two-thirds of the electrons are known to be lost, would certainly lead us to expect the further loss of low energy electrons along the accelerator. The indications are that the additional electrons lost to the beam are those which become decelerated further along the accelerator, and hence become defocused.

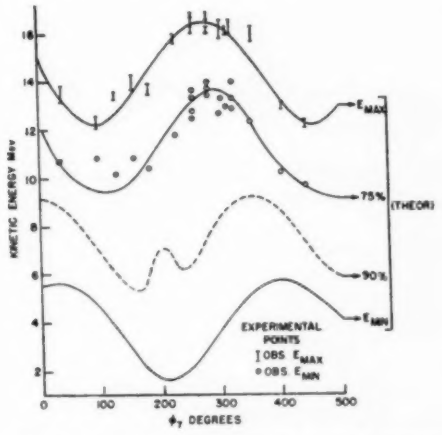


FIG. 10. Effect of variation of phase of seventh section on beam accelerated by seven sections.

BIBLIOGRAPHY

[1] P. T. Demos, "Design and properties of the M.I.T. linear electron accelerator," Technical Report No. 50, L.N.S.E., M.I.T., May 1, 1951.

[2] Halpern, Everhart, Rapuano, and Slater, Phys. Rev. **69**, 688(A) (1946).

[3] R. A. Rapuano and J. Malpern, "Q measurements for high-Q cavities," Technical Report No. 7, R.L.E., M.I.T., June 28, 1946.

[4] L. C. Maier, Jr., "Field strength measurements in resonant cavities," Technical Report No. 143 R.L.E., M.I.T., November 2, 1949.

[5] L. C. Maier, Jr., and J. C. Slater, "Field strength measurements in resonant cavities." J. Appl. Phys. **23**, 68 (1951).

[6] L. C. Maier, Jr., and J. C. Slater, "Determination of axial field strength in a linear accelerator cavity," J. Appl. Phys. **23**, 78 (1951).

[7] J. C. Slater, "The phasing of magnetrons," Technical Report No. 35, R.L.E., M.I.T., April 3, 1947.

[8] J. C. Slater, "The theory of symmetrical wave-guide t's," Technical Report No. 37, R.L.E., M.I.T., April 3, 1947.

[9] J. C. Slater, "Design of linear accelerators," Technical Report No. 47, R.L.E., M.I.T., September 2, 1947.

[10] J. C. Slater, "Electromagnetic waves in iris-loaded wave guides," Technical Report No. 48, R.L.E., M.I.T., September 19, 1947.

[11] J. C. Slater, "The design of linear accelerators," Revs. Modern Phys. **20**, 473 (1948).

[12] J. C. Slater, *Microwave Electronics* (D. Van Nostrand Company, Inc., New York, 1950).

[13] J. R. Terrall, "On the motion of particles in a traveling wave linear accelerator," Phys. Rev. **77**, 415 (1950).

[14] J. R. Terrall and J. C. Slater, "Particle dynamics in the linear accelerator," J. Appl. Phys. **23**, 66 (1951). See also Quarterly Progress Reports of the Research Laboratory of Electronics, M.I.T., to January, 1950, and Quarterly Progress Reports of the Laboratory for Nuclear Science and Engineering, M.I.T., since January, 1950.

Particle Dynamics in the Linear Accelerator*

J. R. TERRALL† AND J. C. SLATER
Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts
(Received August 2, 1951)

Hamilton's equations for the motion of an electron in a linear accelerator are integrated to find the final kinetic energy and phase of an electron injected with arbitrary initial kinetic energy and phase, after traveling down a fixed length of accelerator. The results are expressed in the form of a map of the initial energy-phase space onto a final energy-phase space. This map proves in practice to be very convenient for discussing the actual operation of the accelerator. The curves presented are calculated for numerical values appropriate for the M.I.T. accelerator.

IT has been shown[1] that the longitudinal motion of an electron in a linear accelerator is governed by hamiltonian equations derived from the hamiltonian function

$$H=(m_0^2c^4+p^2c^2)^{\frac{1}{2}}-pv_0-\epsilon m_0c^2\cos\phi. \quad (1)$$

Here p is the relativistic momentum $m_0v/(1-v^2/c^2)^{\frac{1}{2}}$. The term $(m_0^2c^4+p^2c^2)^{\frac{1}{2}}$ is the electron's kinetic energy (including rest energy). The electron is assumed to be moving in a traveling wave whose longitudinal electric field is $E\sin\omega(t-z/v_0)$. The quantity ϵ is an abbreviation for $eEv_0/\omega m_0c^2$. The phase ϕ is an abbreviation for $\omega z'/v_0$, where z' is equal to $z-v_0t$, and measures the electron's position with respect to the traveling wave. It is the purpose of this note to put the dynamics of such electronic motion in a form most useful for discussing the actual operation of the linear accelerator.

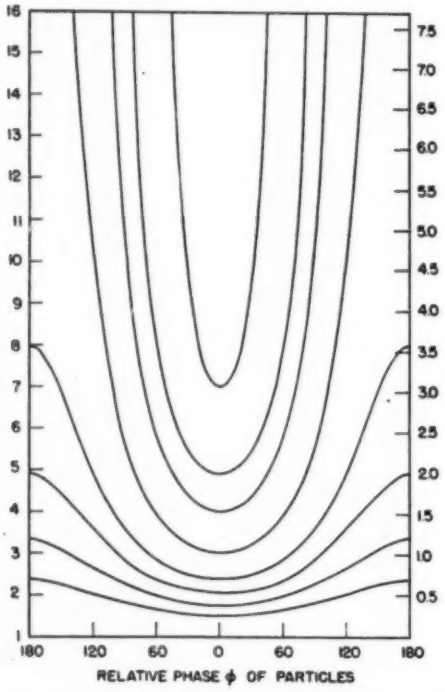


FIG. 1. Lines of constant H in phase space. Ordinates show, left, normalized total energy; right, kinetic energy in Mev of the electrons.

The hamiltonian relation (1) can be conveniently discussed by setting up a phase space, in which the coordinate z' (or the phase ϕ) is the abscissa, and the momentum p the ordinate, and in which the lines of constant H are plotted. Such a plot is given in Fig. 53, (reference 1) for the case $v_0=c/2$, and in Fig. 54 (reference 1) for the case $v_0=c$. We give a similar plot in Fig. 1, which differs from the phase space just described in that we use the kinetic energy of the electron as ordinate rather than the momentum; in the relativistic range these are proportional to each other, but there are differences at low energy. The curves of Fig. 1 are computed for approximately the numerical values appropriate for the M.I.T. linear accelerator, for which $v_0=c$, the resonant wavelength $2\pi c/\omega$ is 10.70 cm, and E is about 24 kev/cm, giving $\epsilon=0.08$. The significance of this figure is then simple: A given electron follows along a line of constant H. Thus, if we know that at an initial instant the electron is found with a given kinetic energy and given phase, we can locate a point on the graph corresponding to this energy and phase, find the contour of constant H passing through that point, and can then read off the subsequent energies and phases which the particle will attain: It will travel in a clockwise direction along a contour, slipping backward in phase as compared to the wave (since it cannot travel with quite the velocity of light), losing energy when its phase is positive, gaining it when its phase is negative. There is a range of injection energies and phases for which the contours eventually rise to infinity, so that the energy of an electron injected in such a phase will eventually increase indefinitely, provided the accelerator is long enough.

This graphical relation gives considerable information, but by no means all we wish for discussing the accelerator. It does not tell us how long it takes for an electron to travel a given distance along a contour of constant energy. We must use Hamilton's equations directly to get this additional information. In the M.I.T. accelerator, electrons are injected all at one energy (actually 2 Mev), and uniformly distributed over phase. We are interested in the energy distribution of the electrons emerging after traversing a fixed length of accelerator. Hamilton's equations, of course, determine

* This work has been supported in part by the Signal Corps, Air Materiel Command, and ONR.
† Now at Tufts College, Medford, Massachusetts.
[1] J. C. Slater, *Microwave Electronics* (D. Van Nostrand Company, Inc., New York, 1950), p. 262.

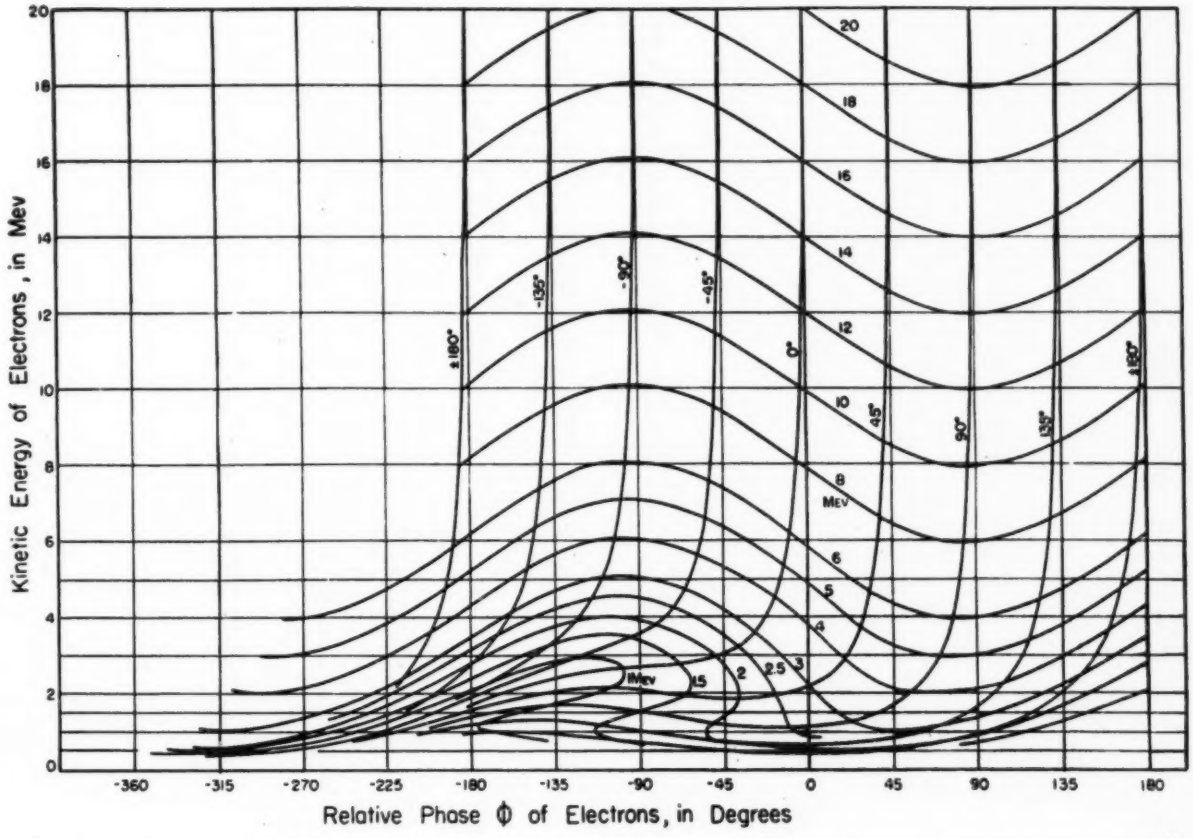


FIG. 2. Map of final energy and phase of electrons after traversing a section of accelerator, in terms of initial energy and phase.

the complete motion of an electron; thus they tell us how rapidly an electron moves around an energy contour in Fig. 1, and hence how far it will move in traversing a given length of accelerator. In other words, if we know the initial energy and phase of an electron on entering a given length of accelerator, Hamilton's equations allow us to find the energy and phase after traversing the length of accelerator. That is, the operation of traveling down the length of accelerator maps any point of the space of Fig. 1 onto another point lying on the same energy contour. Such a map can be exhibited graphically by showing the distorted network of lines into which a rectangular network is mapped by traversing the accelerator. We show such a map in Fig. 2, computed for the same accelerating field as given in Fig. 1, and for an accelerator eight wavelengths long. (This is the length of one of the sections of the M.I.T. accelerator; the complete accelerator consists of seven such sections.) One set of lines of this distorted network is labeled with the initial energy of the electrons, the other set with the initial phase. In the high energy range, we see that the initial phase and final phase almost exactly coincide (since here the electrons are traveling with almost exactly the velocity of light, and cannot slip backward in phase), and the change of energy in the accelerator is almost exactly a sinusoidal function of phase, giving an increase of about 2 Mev at the phase $-90°$ (the maximum accelerating field), and an equal decrease of energy at $+90°$. At the lower energies, the situation is much more complicated; the electrons injected at positive phase angles, which as we see from Fig. 1 initially are slowed down to considerably less than the velocity of light, slip backward considerably in phase, and end up with much less energy than they had on entering.

The diagram of Fig. 2 can be used in a very simple and direct way to discuss the whole behavior of the accelerator. For the first section, injection is at 2 Mev, with equal numbers of electrons in equal intervals of phase. Thus after emerging from this section, electrons will be distributed along the line marked 2 Mev in Fig. 2, with equal numbers in each interval of equal phase, indicated by the contours of constant phase. We see by inspection of the graph that electrons initially injected between phases of $0°$ and $-135°$ will have energies equal to or greater than 3 Mev after acceleration, with some electrons having as much as 4 Mev. Furthermore, these electrons will be distributed in final phase from about $-45°$ to $-180°$. We can now take the distribution as it emerges from this section, take the energy and phase of any one of its electrons as the initial conditions for the next section, and follow the progress of the electrons and their distribution through the second section. This can be continued for

all seven sections. Such an analysis of the behavior of the electron beam, with experimental verification, is presented in the accompanying paper by P. T. Demos, A. F. Kip, and J. C. Slater.[2]

The calculation of Fig. 2 was made as follows. We start with the equation $z'=z-ct$, and differentiate to get

$$dz'=dz-cdt. \tag{2}$$

Furthermore, from Hamilton's equations, we have

$$dz'/dt=\partial H/\partial p=pc^2(m_0^2c^4+p^2c^2)^{-\frac{1}{2}}-c.$$

This allows us to write dt in terms of z'. We do this, substitute in Eq. (2), and find

$$dz=dz'\left[\frac{pc}{pc-(m_0^2c^4+p^2c^2)^{\frac{1}{2}}}\right]=-dz'\left[\frac{pc}{H+m_0c^2\epsilon\cos\phi}\right], \tag{3}$$

where we have made use of Eq. (1). We can now solve (1) for pc, and find

$$pc=[m_0^2c^4-(H+m_0c^2\epsilon\cos\phi)^2]/2(H+m_0c^2\epsilon\cos\phi). \tag{4}$$

[2] Demos, Kip, and Slater, J. Appl. Phys. **23**, 53 (1951).

We substitute (4) in (3), set $\phi=\omega z'/v_0$, and have

$$(2\omega/c)dz=[1-(H/m_0c^2+\epsilon\cos\phi)-2]d\phi. \tag{5}$$

This equation can be integrated to give the distance z traveled by the electron, as a function of the final phase ϕ and initial phase ϕ_0. The integration can be performed analytically, but the resulting expression is so complicated that it has proved simpler in practice to integrate numerically. When the relation between ϕ and z is thus established, for each H value, we need only substitute in (1) to find the kinetic energy $(m_0^2c^4+p^2c^2)^{\frac{1}{2}}$ after traversing the required distance. Thus we know the final phase and energy in terms of the initial phase and energy, and are able to construct the curves of Fig. 2.

We are indebted to the Computing Group associated with the M.I.T. Laboratory for Nuclear Science and Engineering and the Research Laboratory of Electronics, consisting of Pauline Leighton, Patricia Boland, Elizabeth Campbell, Evelyn Hill, Mida Karakashian, Barbara Levine, and Hannah Paul, for their help in carrying out the computations.

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 1 JANUARY, 1952

Field Strength Measurements in Resonant Cavities*

L. C. MAIER, JR.† AND J. C. SLATER
Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts

(Received August 2, 1951)

It is shown that the frequency of a resonant electromagnetic cavity is perturbed by inserting a metallic sphere, needle, or disk of dimensions small compared to a wavelength by an amount depending upon the local electric and magnetic field at the position of the perturbing object. This perturbation is calculated for ellipsoidal objects of needle-shaped, spherical, and disk-shaped form. The perturbations by the different objects depend upon different components of electric and magnetic fields, and by combining measurements with all three, it is in theory possible to measure all the field components. Experimental checks of the calculations are described, resulting in satisfactory agreement between theory and experiment except with the needles, in which the perturbation is very sensitive to the precise shape of the object, and the needles used were not accurate enough ellipsoids to give satisfactorily quantitative results.

IN a resonant electromagnetic cavity of complicated shape, it may be nearly impossible to calculate the distribution of electric and magnetic fields, and an experimental method of measuring these fields becomes highly desirable. Probes or loops extending in from the walls, and feeding power out to external coaxial lines and power measuring devices, furnish a familiar method of measuring the field distribution around the walls of the cavity, but cannot be used in the interior. A quite different approach to the problem is provided by the theorem giving the change of the resonant frequency of a cavity by a perturbation of its boundaries. This theorem may be stated in the form,[1]

$$\omega^2=\omega_0^2\left(1+\int(H_a^2-E_a^2)dv\right), \tag{1}$$

where ω_0 is the unperturbed angular frequency of a cavity, ω its perturbed frequency. The perturbation is assumed to be produced by pushing in a section of the wall of the cavity by an infinitesimal amount. H_a and E_a are proportional to the magnetic and electric fields in the unperturbed cavity, normalized so that the integral of H_a^2 or E_a^2 over the cavity is unity. The integral in (1) is to be taken over the volume removed from the cavity by pushing in the walls. We see, then,

* This work has been supported in part by the Signal Corps, the Air Materiel Command, and ONR.

† Now at Electronics Laboratory, General Electric Company, Syracuse, New York.

[1] J. C. Slater, *Microwave Electronics* (D. Van Nostrand Company, Inc., New York, 1950), p. 81.

that if the wall is pushed in by a small amount in the neighborhood of a point, the frequency will be perturbed by an amount depending upon the value of $H_a^2-E_a^2$ at that point, so that by measurement of the frequency change, we can evaluate this quantity. A method developed by Hansen and Post[2] makes use of essentially this principle.

When we consider the theorem (1), it seems reasonable to suppose that a conducting obstacle, such as a small metallic sphere, introduced into the interior of a cavity (suspended, for instance, by a thread), should perturb the frequency by an amount depending upon the local electric and magnetic fields, and thus that we could use this perturbation to measure the field strength at an interior point. It is the object of the present paper to show that this is the case, and to evaluate the amount of the perturbation. We might further suppose that different shaped obstacles would have different effects, and would react differently to electric and magnetic fields, so that we might have methods of separately evaluating both types of field, and of finding their directions. Harries[3] has used a thin metallic disk in a cavity to find the direction of the electric field. He notes that if the plane of the disk is normal to the electric field, the electric lines of force will terminate on both faces of the disk, and will have the same form as if they went right through the disk, so that the disk in this position will not perturb the field, and hence will not affect the frequency. On the other hand, if the disk is rotated so that the lines of force lie in the plane of the disk, there will be a large perturbation. Just the opposite case is found with the disk in a magnetic field: if it is normal to the field, the lines of magnetic force will be greatly perturbed, while if the magnetic field is parallel to the plane of the disk the tangential magnetic lines of force will not be affected. Another geometrical shape which is useful is a needle-shaped rod or wire. If it is placed parallel to the electric lines of force, it will attract the lines like a lightning rod, and will cause a large perturbation, while if it is at right angles it will cause small perturbation.

All the geometrical shapes considered in the preceding paragraph can be approximated by special cases of ellipsoids of revolution. Therefore we have considered the problem of the perturbation of frequency of an arbitrary resonant cavity, with a small conducting ellipsoid of revolution inserted at an arbitrary point. The problem can be solved exactly, provided the dimensions of the ellipsoid are sufficiently small compared to a wavelength. The reason is that in this case the field near the ellipsoid reduces with sufficient accuracy to a static field, for which the solutions can be calculated. It might be thought at first sight that all we had to do to solve the problem would be to integrate the quantity $H_a^2-E_a^2$ of (1) over the volume of the ellipsoid, so that we might think that we merely had to replace the integral by the local value of $H_a^2-E_a^2$ in the absence of the ellipsoid, times the volume of the ellipsoid. This is far from the case; the ellipsoid perturbs the field in a way which is essential to the problem. We can, however, use (1) to find the additional change in frequency if the size of the ellipsoid is slightly altered, the integration being performed over the shell between the original and infinitesimally perturbed ellipsoids. The resulting formula for change of frequency can then be integrated as the ellipsoid expands from zero dimensions to the actual size, and in this way we get the desired results.

We shall not give the complete analysis of the problem in the present paper; it is presented elsewhere.[4] We shall carry the case of the sphere through completely, and this illustrates the nature of the method; then we shall outline the procedure and state the results for the ellipsoids of revolution. Finally we shall describe experimental tests which have been made to check the correctness of the theory. The method has been developed into a practical scheme for measuring field strength in the linear accelerator cavity, and its application to this problem is described in another paper.[5]

1. FIELD PERTURBATION BY A SPHERE

We shall first consider the case of an electric field E_0 directed along the positive z axis, and shall ask for the modified field when a small perfectly conducting metallic sphere of radius r_0 is placed in this field. We know by elementary electrostatics that the electrostatic potential at points outside the sphere is

$$\phi=-E_0 r\cos\theta+(E_0 r_0^3/r^2)\cos\theta. \tag{2}$$

The resulting field components are obtained from the gradient of this potential and are

$$E_r=-(\partial\phi/\partial r)=E_0[1+(2r_0^3/r^3)]\cos\theta \tag{3}$$

$$E_\theta=-(1/r)(\partial\phi/\partial\theta)=E_0[-1+(r_0^3/r^3)]\sin\theta \tag{4}$$

$$E_\psi=-1/(r\sin\theta)(\partial\phi/\partial\psi)=0. \tag{5}$$

The perturbation formula for zero H_a field is

$$[(\omega_0^2-\omega^2)/\omega_0^2]=\int_{\Delta V}E_a^2 dv. \tag{6}$$

Since

$$E_a^2=E_r^2+E_\theta^2+E_\psi^2 \tag{7}$$

and

$$dv=r^2\sin\theta dr d\theta d\psi \tag{8}$$

integrating the field for an infinitesimal expansion of

[2] W. W. Hansen and R. W. Post, J. Appl. Phys. **19**, 1059 (1948).
[3] J. H. O. Harries, Wireless Engineer **24**, 135 (1937).
[4] L. C. Maier, Jr., Technical Report No. 143, Research Laboratory of Electronics, Massachusetts Institute of Technology, November 2, 1949.
[5] L. C. Maier, Jr., and J. C. Slater, "Determination of field strength in a linear accelerator cavity," J. Appl. Phys. **23**, 78 (1952).

the sphere into the field, from r_0 to r_0+dr_0, gives

$$\left(\int_{r_0}^{r_0+dr_0}[1+(2r_0^3/r^3)]^2r^2dr\int_0^\pi\cos^2\theta\sin\theta d\theta+\int_{r_0}^{r_0+dr_0}[1-(r_0^3/r^3)]^2r^2dr\int_0^\pi\sin^3\theta d\theta\right)2\pi E_0^2=\int_V^{V+dV}E_a^2dv. \quad (9)$$

It is obvious, however, that

$$\int_{x_0}^{x_0+dx_0}f(x)dx=f(x_0)dx_0 \quad (10)$$

and applying this to the previous expression, the result for the infinitesimal expansion of the sphere is

$$\int_V^{V+dV}E_a^2dv=12\pi E_0^2r_0^2dr_0. \quad (11)$$

For the total perturbation caused by a sphere of radius r_0, the expression given in (11) must be integrated from zero to r_0, and substituting the resulting expression into the perturbation formula gives

$$[(\omega_0^2-\omega^2)/\omega_0^2]=3E_0^2(4\pi/3)r_0^3. \quad (12)$$

This, then, gives an expression relating the magnitude E_0 of the electric field to the size and shape of a perturbing volume, and the frequency shift caused by its introduction into the field.

The application of the perturbation formula to a metallic sphere in the magnetic field is carried out in a completely analogous manner, and results in

$$[(\omega_0^2-\omega^2)/\omega_0^2]=-(3/2)H_0^2(4\pi/3)r_0^3. \quad (13)$$

By applying the superposition principle, the perturbation in a region where both fields are present is given by

$$[(\omega_0^2-\omega^2)/\omega_0^2]=(4\pi/3)r_0^3(E_0^2-\tfrac{1}{2}H_0^2). \quad (14)$$

From Eq. (14), it is clear that a sphere cannot measure either electric or magnetic field separately, but only the combination $E_a^2-\frac{1}{2}H_a^2$. Often, however, we are interested in one type of field at a point where we know independently that the other type is zero, as for instance along the axis of the linear accelerator, where the magnetic field must be zero by symmetry. In such a case the sphere provides complete information.

2. PERTURBATIONS BY CIRCULAR METALLIC NEEDLES

The next two sections will be concerned with the perturbation effects of ellipsoids of revolution; the present section with ellipses rotated around the major axis to form circular needles, and the following section with ellipses rotated around the minor axis to form circular disks. These two sections will give the necessary additional relations for complete determination of the fields.

The procedure used in finding the potential is that of generalized coordinates and is in the literature.[6] The coordinates to be used for the circular needle of elliptic cross section are the points in space determined by the intersections of a family of confocal ellipsoids, ξ equals constant, with a family of confocal hyperbolas, η equals constant, and with an angle, ψ equals constant, measured around the x axis. The equations of the coordinate surfaces are given by

$$\begin{aligned}&[x^2/(\xi+a^2)]+[r^2/(\xi+c^2)]=1,\\&[x^2/(a^2-\eta)]-[r^2/(\eta-c^2)]=1\end{aligned} \quad (15)$$

where

$$\begin{aligned}&a\geqslant c\\&\xi>-c^2\\&c^2<\eta<a^2\\&r^2=y^2+z^2\\&y=r\sin\psi\\&z=r\cos\psi.\end{aligned} \quad (16)$$

These equations can be solved simultaneously for x, y, and z, which results in

$$x=\left[\frac{(\xi+a^2)(a^2-\eta)}{(a^2-c^2)}\right]^{\frac{1}{2}} \quad (17)$$

$$y=\left[\frac{(\xi+c^2)(\eta-c^2)}{(a^2-c^2)}\right]^{\frac{1}{2}}\sin\psi \quad (18)$$

$$z=\left[\frac{(\xi+c^2)(\eta-c^2)}{(a^2-c^2)}\right]^{\frac{1}{2}}\cos\psi. \quad (19)$$

Since the desired coordinates are ξ, η, and ψ, the metrical coefficients can be calculated from (17), (18), and (19), giving

$$h_1=\frac{1}{2}\left[\frac{(\xi+\eta)}{(\xi+a^2)(\xi+c^2)}\right]^{\frac{1}{2}}\quad h_2=\frac{1}{2}\left[\frac{(\xi+\eta)}{(a^2-\eta)(\eta-c^2)}\right]^{\frac{1}{2}}$$

$$h_3=\left[\frac{(\xi+c^2)(\eta-c^2)}{(a^2-c^2)}\right]^{\frac{1}{2}}. \quad (20)$$

The potential function of the field containing the ellipsoid must be calculated.[7] For the case of a constant field of magnitude C_0, directed along the x axis or along the major axis of the ellipsoid, the original constant field potential is given by

$$\phi_0=-C_0x=-C_0\left[\frac{(\xi+a^2)(a^2-\eta)}{(a^2-c^2)}\right]^{\frac{1}{2}}$$

$$\phi_0=A_1F_1(\xi)F_2(\eta)F_3(\psi). \quad (21)$$

[6] J. A. Stratton, *Electromagnetic Theory* (McGraw-Hill Book Company, Inc., New York, 1941), pp. 48–51.

[7] See reference 6, pp. 209–211.

The desired potential function must satisfy Laplace's equation; it must reduce to a constant on the surface ξ equals zero, that is, on the surface of the needle; and it must be regular at infinity. The primary field potential ϕ_0 is a solution of Laplace's equation, but it is not regular at infinity.

If the boundary surfaces are to be satisfied on the surface ξ equals zero, then ϕ_1, the potential of the induced charge distribution must vary identically with ϕ_0 over surfaces of constant ξ. Assume then, that

$$\phi_1 = A_2 G_1(\xi) F_2(\eta) F_3(\psi). \tag{22}$$

ϕ_1 must also satisfy Lapace's equation, which, in the present coordinate system, is given by

$$\nabla^2\phi = 0 = (\xi+a^2)^{\frac{1}{2}}(\partial/\partial\xi)[(\xi+a^2)^{\frac{1}{2}}(\xi+c^2)(\partial\phi/\partial\xi)] + (a^2-\eta)^{\frac{1}{2}}\frac{\partial}{\partial\eta}\left[(a^2-\eta)^{\frac{1}{2}}(\eta-c^2)\frac{\partial\phi}{\partial\eta}\right] + \frac{1}{4}\left[\frac{(\xi+\eta)(a^2-c^2)}{(\xi+c^2)(\eta-c^2)}\right]\frac{\partial^2\phi}{\partial\psi^2}. \tag{23}$$

Substitution of the expression for ϕ_1 into (23) gives the following differential equation for $G_1(\xi)$

$$\frac{d^2G_1}{d\xi^2} + \frac{dG_1}{d\xi}\left[\frac{1}{2}\frac{1}{\xi+a^2} + \frac{1}{\xi+c^2}\right] - \frac{G_1}{2}\frac{1}{(\xi+a^2)(\xi+c^2)} = 0. \tag{24}$$

If one solution of a second-order, linear differential equation is known, in this case, $F_1(\xi)$, an independent solution, $G_1(\xi)$, can be determined from it by integration.[8] That is, if $y_1(x)$ is a solution of

$$(d^2y/dx^2) + p(x)(dy/dx) + q(x)y = 0 \tag{25}$$

then

$$y_2(x) = y_1(x)\int\frac{\exp[-\int p(x)dx]}{[y_1(x)]^2}dx. \tag{26}$$

Applying this to find $G_1(\xi)$ where $F_1(\xi)$ is given from (21) and (17), the result is that

$$G_1(\xi) = (\xi+a^2)^{\frac{1}{2}}\int\frac{\exp\{-\frac{1}{2}\int[d\xi/(\xi+a^2)]\}\exp\{-\int[d\xi/(\xi+c^2)]\}}{(\xi+a^2)}d\xi \tag{27}$$

$$G_1(\xi) = (\xi+a^2)^{\frac{1}{2}}\int\frac{d\xi}{(\xi+a^2)^{\frac{3}{2}}(\xi+c^2)}. \tag{28}$$

The limits of integration on $G_1(\xi)$ are arbitrary, but to make ϕ_1 vanish at infinity, the limits used are ξ to infinity. Substituting (27) into (22) gives

$$\phi_1 = \phi_0\frac{A_2}{A_1}\int_\xi^\infty\frac{d\xi}{(\xi+a^2)^{\frac{1}{2}}(\xi+c^2)}. \tag{29}$$

The total potential for this case of the applied field parallel to the major axis of the needle is the sum of primary potential and the induced potential, and is

$$\phi_{11} = -C_0\left[\frac{(\xi+a^2)(a^2-\eta)}{(a^2-c^2)}\right]^{\frac{1}{2}} \times\left[1+A\int_\xi^\infty\frac{d\xi}{(\xi+a^2)^{\frac{1}{2}}(\xi+c^2)}\right]. \tag{30}$$

Since nothing in the foregoing derivation restricted it to either an electric or magnetic field, the constant C_0 can be either E_0 or H_0, and A will then be determined from the boundary conditions applying to the field being considered.

For the case of the primary field being directed along the y or z axis, it is simple to sketch the derivation of the potential to find the changes in the final function. $F_1(\xi)$ will in this case be $(\xi+C^2)^{\frac{1}{2}}$ and the resulting integration for $G_1(\xi)$ will give

$$G_1(\xi) = (\xi+c^2)^{\frac{1}{2}}\int_\xi^\infty\frac{d\xi}{(\xi+c^2)^2(\xi+a^2)^{\frac{1}{2}}}. \tag{31}$$

The final potential for the primary field perpendicular to the major axis will be

$$\phi_{\perp} = -C_0\left[\frac{(\xi+c^2)(\eta-c^2)}{(a^2-c^2)}\right]^{\frac{1}{2}} \times\sin\psi\left[1+A\int_\xi^\infty\frac{d\xi}{(\xi+c^2)^2(\xi+a^2)^{\frac{1}{2}}}\right]. \tag{32}$$

There are obviously four cases which must be calculated, two for E_0, parallel and perpendicular to the major axis of the needle, and two for H_0. The procedure is similar to that followed for the sphere. The case of E_0 parallel to the major axis will be calculated in detail to illustrate the method, but only the final results of the other three possible orientations will be given. Consider a needle of semi-major axis a and semi-minor axis c placed in a constant electric field of magnitude E_0. The potential in the presence of the needle, as found in (30), is given by

$$\phi_{11} = -E_0\left[\frac{(\xi+a^2)(a^2-\eta)}{(a^2-c^2)}\right]^{\frac{1}{2}}[1+A_E V] \tag{33}$$

[8] E. L. Ince, *Ordinary Differential Equations* (Longmans, Green, and Company, London, 1927), p. 122.

where

$$V=\int_{\xi}^{\infty}\frac{d\xi}{(\xi+a^2)^{\frac{1}{2}}(\xi+c^2)} \tag{34}$$

$$=\frac{-1}{(a^2-c^2)^{\frac{1}{2}}}\left[\frac{(a^2-c^2)^{\frac{1}{2}}}{(\xi+a^2)^{\frac{1}{2}}}+\ln\frac{(\xi+a^2)^{\frac{1}{2}}-(a^2-c^2)^{\frac{1}{2}}}{(\xi+a^2)^{\frac{1}{2}}+(a^2-c^2)^{\frac{1}{2}}}\right]. \tag{35}$$

The field components, E_ξ, E_η, and E_ψ are found from the gradient of this potential, and the value of A_E from the boundary condition that the tangential components of the electric field, E_η and E_ψ, are zero on the surface of the needle, ξ equals zero. Since

$$E_\eta=-\frac{1}{h_2}\frac{\partial\phi_{11}}{\partial\eta}=E_0\left[\frac{(\xi+a^2)(\eta-c^2)}{(\xi+\eta)(a^2-c^2)}\right]^{\frac{1}{2}}[1+A_EV] \tag{36}$$

and if E_η equals zero at ξ equals zero for all values of η and ψ, then

$$A_E=(-1/V_0) \tag{37}$$

where

$$V_0=\frac{-1}{(a^2-c^2)^{\frac{1}{2}}}\left[\frac{2(a^2-c^2)^{\frac{1}{2}}}{a}+\ln\frac{a-(a^2-c^2)^{\frac{1}{2}}}{a+(a^2-c^2)^{\frac{1}{2}}}\right]. \tag{38}$$

This makes

$$E_\eta=E_0\left[\frac{(\xi+a^2)(\eta-c^2)}{(\xi+\eta)(a^2-c^2)}\right]^{\frac{1}{2}}\left[1-\frac{V}{V_0}\right] \tag{39}$$

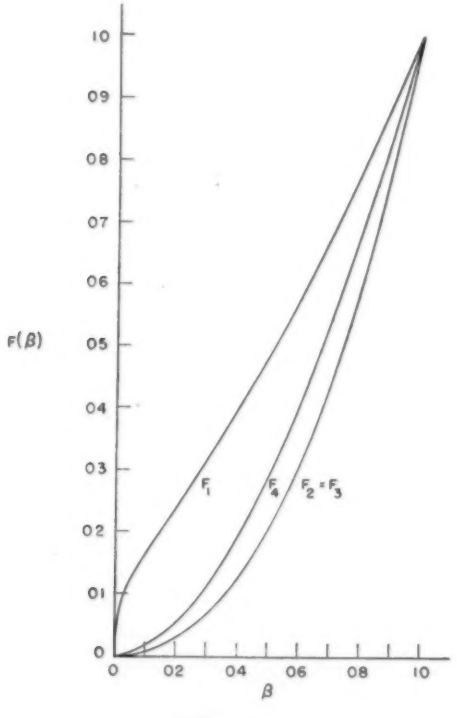


FIG. 1.

$$E_\xi=E_0\left[\frac{(\xi+c^2)(a^2-\eta)}{(\xi+\eta)(a^2-c^2)}\right]^{\frac{1}{2}}\times\left[1-\frac{V}{V_0}+\frac{2}{V_0(\xi+a^2)^{\frac{1}{2}}(\xi+c^2)}\right] \tag{40}$$

$$E_\psi=0. \tag{41}$$

The next step is to apply the perturbation formula to an expansion of the needle, and integrate the field over the expansion. The perturbation formula, however, holds only for an expansion of the perturbing volume to a similar volume. In the case of the sphere this was automatic, since any two spheres are similar, but here the needle can be expanded in many ways. The surface ξ equals zero represents the surface of the needle, and it might be thought that an expansion from ξ equals zero to $d\xi$, or from one ellipsoid to a confocal ellipsoid infinitesimally larger would be wanted. The perturbation formula would not hold in this case, since the needle would not be expanding all dimensions equally into the field, as is necessary for correct application. What is then needed is the equation of a needle similar to the original, but infinitesimally larger. Since the equation of the basic needle is

$$(x^2/a^2)+(r^2/c^2)=1 \tag{42}$$

the equation for a similar needle is

$$(x^2/a^2)+(r^2/c^2)=\alpha^2 \tag{43}$$

where

$$\alpha^2=(a'^2/a^2)=(c'^2/c^2).$$

Putting the values of x^2 and r^2 found in (17), (18), and (19) into the expression for the similar needle, and solving for the value of ξ_0 denoting the surface of the similar needle, gives

$$\xi_0=\{[a^2c^2(\alpha^2-1)]/\eta\}. \tag{44}$$

If the similar needle is only infinitesimally greater than the basic one, then

$$\alpha=(a'/a)=[(a+da)/a]=(c'/c)=[(c+dc)/c]=1+d\alpha \tag{45}$$

and

$$d\xi_0=[(2a^2c^2d\alpha)/\eta]. \tag{46}$$

This value will be the upper limit of integration for the application of the perturbation formula to the expansion of the basic ellipsoid. The volume increment is given by

$$dv=\frac{1}{2}\frac{(\xi+\eta)d\xi d\eta d\psi}{(\xi+a^2)^{\frac{1}{2}}(a^2-\eta)^{\frac{1}{2}}(a^2-c^2)^{\frac{1}{2}}}. \tag{47}$$

With this information, the components E_ξ, E_η, and E_ψ found in (38), (39), and (40) can now be integrated over

the infinitesimal expansion of the needle. This results in

$$\int_V^{V+dV} E^2 dv = \frac{1}{2}\int_0^{2\pi} d\psi \int_{c^2}^{a^2} d\eta \int_0^{d\xi_0} \times d\xi (E_\xi^2 + E_\eta^2 + E_\psi^2)\left[\frac{(\xi+\eta)}{(\xi+a^2)^{\frac{1}{2}}(a^2-\eta)^{\frac{1}{2}}(a^2-c^2)^{\frac{1}{2}}}\right]. \quad (48)$$

However, E_ψ equals zero, and E_ξ and E_η both contain terms in $(1-V/V_0)$, which, when integrated between the limits on ξ, are equal to zero. Carrying out the remaining integrations over ξ gives

$$\int_V^{V+dV} E^2 dv = \frac{\pi E_0^2}{(a^2-c^2)^{\frac{1}{2}}} \frac{c^2}{a} \frac{4}{V_0^2 a^2 c^4} \int_c^{a^2} (a^2-\eta)^{\frac{1}{2}} d\eta d\xi_0. \quad (49)$$

Substituting the value of $d\xi_0$ given in (45), and integrating over η gives

$$\int_V^{V+dV} E^2 dv = \frac{-8\pi E_0^2}{(a^2-c^2)^{\frac{1}{2}}} \frac{d\alpha}{V_0^2}\left[\frac{2(a^2-c^2)^{\frac{1}{2}}}{a} + \ln\frac{a-(a^2-c^2)^{\frac{1}{2}}}{a+(a^2-c^2)^{\frac{1}{2}}}\right] \quad (50)$$

and substituting the value of V_0 from (37) gives

$$\int_V^{V+dV} E^2 dv = \frac{8\pi E_0^2 (a^2-c^2)^{\frac{1}{2}} d\alpha}{\left[\ln\frac{a+(a^2-c^2)^{\frac{1}{2}}}{a-(a^2-c^2)^{\frac{1}{2}}} - \frac{2(a^2-c^2)^{\frac{1}{2}}}{a}\right]}. \quad (51)$$

This is the result of the integration over the infinitesimal expansion of the needle. To find the total perturbation caused by the introduction of the needle, the foregoing expression must be integrated from zero to a and c. For similar needles, β equals a/c is a constant and $d\alpha$ equals $(da)/a$ equals $(dc)/c$. Substituting this into (50) and integrating gives the total perturbation caused by the introduction of the needle into the previously constant field

$$\frac{\omega_0^2-\omega^2}{\omega_0^2} = \frac{8\pi E_0^2(1-\beta^2)^{\frac{1}{2}} \int_0^a a^2 da}{\left[\ln\frac{1+(1-\beta^2)^{\frac{1}{2}}}{1-(1-\beta^2)^{\frac{1}{2}}} - 2(1-\beta^2)^{\frac{1}{2}}\right]} = \frac{2E_0^2(1-\beta^2)^{\frac{1}{2}}(4\pi/3)a^3}{\left[\ln\frac{1+(1-\beta^2)^{\frac{1}{2}}}{1-(1-\beta^2)^{\frac{1}{2}}} - 2(1-\beta^2)^{\frac{1}{2}}\right]}. \quad (52)$$

This is an exact expression for perturbation produced by a needle of semi-major axis a and semi-minor axis c. It holds for all values of β; and in particular, as β approaches unity, a approaches c, the needle approaches a sphere, and the expression reduces to

$$[(\omega_0^2-\omega^2)/\omega_0^2] = 3E_0^2(4\pi/3)a^3 \quad (53)$$

which is the value found in (12) for the perturbation of a metallic sphere. Figure 1 contains the foregoing relation normalized to the perturbation effect of a sphere of radius r_0 equal to a, and plotted as $F_1(\beta)$ from zero to one. Figure 2 contains the same function on an expanded scale for small β.

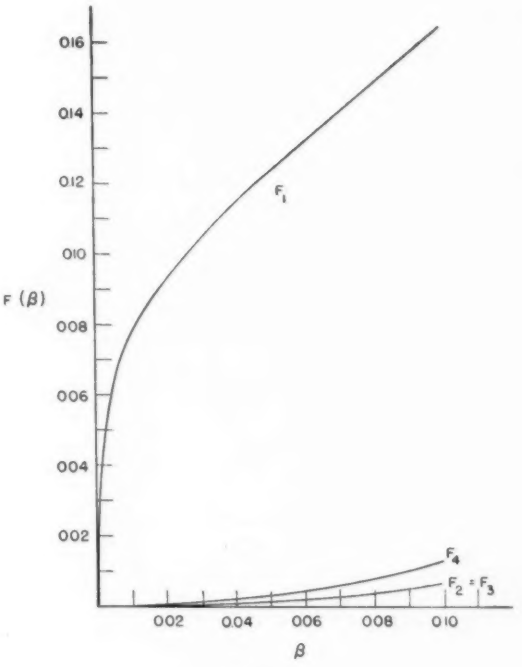


FIG. 2.

The same procedure is followed for the applied E field perpendicular to the major axis of the needle, and for the two orientations in the H field; except, of course, for using the correct potential and the correct boundary conditions for the specific case being considered. The results of these, as well as the foregoing, are summarized as follows, and are also plotted in Figs. 1 and 2.

E_0 parallel to a,

$$\frac{\omega_0^2-\omega^2}{\omega_0^2} = \frac{E_0^2(1-\beta^2)^{\frac{1}{2}}(4\pi/3)a^3}{\left[\frac{1}{2}\ln\frac{1+(1-\beta^2)^{\frac{1}{2}}}{1-(1-\beta^2)^{\frac{1}{2}}} - (1-\beta^2)^{\frac{1}{2}}\right]} \quad (54)$$

$$F_1(\beta) = \frac{\frac{1}{3}(1-\beta^2)^{\frac{1}{2}}}{\left[\frac{1}{2}\ln\frac{1+(1-\beta^2)^{\frac{1}{2}}}{1-(1-\beta^2)^{\frac{1}{2}}} - (1-\beta^2)^{\frac{1}{2}}\right]}. \quad (55)$$

E_0 perpendicular to a,

$$\frac{\omega_0^2-\omega^2}{\omega_0^2} = \frac{2E_0^2(1-\beta^2)^{\frac{1}{2}}(4\pi/3)a^3}{\left[\frac{(1-\beta^2)^{\frac{1}{2}}}{\beta^2} - \frac{1}{2}\ln\frac{1+(1-\beta^2)^{\frac{1}{2}}}{1-(1-\beta^2)^{\frac{1}{2}}}\right]} \quad (56)$$

$$F_2(\beta) = \frac{\frac{2}{3}(1-\beta^2)^{\frac{1}{2}}}{\left[\frac{(1-\beta^2)^{\frac{1}{2}}}{\beta^2} - \frac{1}{2}\ln\frac{1+(1-\beta^2)^{\frac{1}{2}}}{1-(1-\beta^2)^{\frac{1}{2}}}\right]}. \quad (57)$$

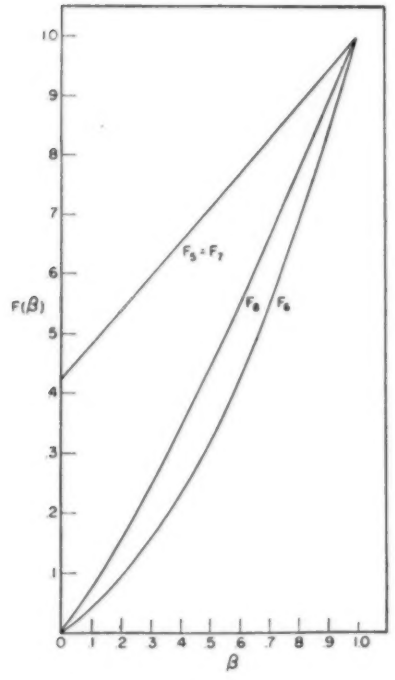


FIG. 3.

H_0 parallel to a,

$$\frac{\omega_0^2-\omega^2}{\omega_0^2}=\frac{-H_0^2(1-\beta^2)^{\frac{1}{2}}(4\pi/3)a^3}{\left[\frac{(1-\beta^2)^{\frac{1}{2}}}{\beta^2}-\frac{1}{2}\ln\frac{1+(1-\beta^2)^{\frac{1}{2}}}{1-(1-\beta^2)^{\frac{1}{2}}}\right]} \tag{58}$$

$$F_3(\beta)=\frac{\frac{2}{3}(1-\beta^2)^{\frac{1}{2}}}{\left[\frac{(1-\beta^2)^{\frac{1}{2}}}{\beta^2}-\frac{1}{2}\ln\frac{1+(1-\beta^2)^{\frac{1}{2}}}{1-(1-\beta^2)^{\frac{1}{2}}}\right]}. \tag{59}$$

H_0 perpendicular to a,

$$\frac{\omega_0^2-\omega^2}{\omega_0^2}=-\frac{2H_0^2(1-\beta^2)^{\frac{1}{2}}(4\pi/3)a^3}{\left[\frac{1-2\beta^2}{\beta^2}(1-\beta^2)^{\frac{1}{2}}+\frac{1}{2}\ln\frac{1+(1-\beta^2)^{\frac{1}{2}}}{1-(1-\beta^2)^{\frac{1}{2}}}\right]} \tag{60}$$

$$F_4(\beta)=\frac{(4/3)(1-\beta^2)^{\frac{1}{2}}}{\left[\frac{1-2\beta^2}{\beta^2}(1-\beta^2)^{\frac{1}{2}}+\frac{1}{2}\ln\frac{1+(1-\beta^2)^{\frac{1}{2}}}{1-(1-\beta^2)^{\frac{1}{2}}}\right]}. \tag{61}$$

The $F(\beta)$ functions are the ratios of the frequency shift of a needle of semi-major axis a and semi-minor axis c to the frequency shift of a sphere of radius a placed in the same field. The method of applying these results is that the calculated perturbation of a sphere of radius a is multiplied by the value of $F(\beta)$ corresponding to the dimensions of the needle and its orientation with respect to the field. This gives the desired relation between the shape of the perturbation, its size and orientation, and the field strength.

3. PERTURBATIONS BY CIRCULAR METALLIC DISKS

The perturbations by circular metallic disks, formed by rotating ellipses around their semi-minor axis, can be found following the procedure given in the previous section. In this case it is convenient to introduce a new coordinate system defined by

$$\begin{aligned}&[r^2/(\xi+a^2)]+[z^2/(\xi+c^2)]=1,\\&[r^2/(a^2-\eta)]-[z^2/(\eta-c^2)]=1.\end{aligned} \tag{62}$$

Using these coordinates and following an analagous procedure gives the final results which follow in both frequency perturbation form and normalized to the perturbation of a sphere of radius equal to the semi-major axis of the disk.

E_0 parallel to a,

$$\frac{\omega_0^2-\omega^2}{\omega_0^2}=\frac{2E_0^2(1-\beta^2)^{\frac{1}{2}}(4\pi/3)a^3}{\left[\frac{\pi}{2}-\tan^{-1}\frac{\beta}{(1-\beta^2)^{\frac{1}{2}}}-\beta(1-\beta^2)^{\frac{1}{2}}\right]} \tag{63}$$

$$F_5(\beta)=\frac{\frac{2}{3}(1-\beta^2)^{\frac{1}{2}}}{\left[\frac{\pi}{2}-\tan^{-1}\frac{\beta}{(1-\beta^2)^{\frac{1}{2}}}-\beta(1-\beta^2)^{\frac{1}{2}}\right]}. \tag{64}$$

E_0 perpendicular to a,

$$\frac{\omega_0^2-\omega^2}{\omega_0^2}=\frac{E_0^2(1-\beta^2)^{\frac{1}{2}}(4\pi/3)a^3}{\left[\frac{(1-\beta^2)^{\frac{1}{2}}}{\beta}+\tan^{-1}\frac{\beta}{(1-\beta^2)^{\frac{1}{2}}}-\frac{\pi}{2}\right]} \tag{65}$$

$$F_6(\beta)=\frac{\frac{1}{3}(1-\beta^2)^{\frac{1}{2}}}{\left[\frac{(1-\beta^2)^{\frac{1}{2}}}{\beta}+\tan^{-1}\frac{\beta}{(1-\beta^2)^{\frac{1}{2}}}-\frac{\pi}{2}\right]}. \tag{66}$$

H_0 perpendicular to a,

$$\frac{\omega_0^2-\omega^2}{\omega_0^2}=\frac{-H_0^2(1-\beta^2)^{\frac{1}{2}}(4\pi/3)a^3}{\left[\frac{\pi}{2}-\tan^{-1}\frac{\beta}{(1-\beta^2)^{\frac{1}{2}}}-\beta(1-\beta^2)^{\frac{1}{2}}\right]} \tag{67}$$

$$F_7(\beta)=\frac{\frac{2}{3}(1-\beta^2)^{\frac{1}{2}}}{\left[\frac{\pi}{2}-\tan^{-1}\frac{\beta}{(1-\beta^2)^{\frac{1}{2}}}-\beta(1-\beta^2)^{\frac{1}{2}}\right]}. \tag{68}$$

H_0 parallel to a,

$$\frac{\omega_0^2-\omega^2}{\omega_0^2}=\frac{-2H_0^2(1-\beta^2)^{\frac{1}{2}}(4\pi/3)a^3}{\left[\tan^{-1}\frac{\beta}{(1-\beta^2)^{\frac{1}{2}}}+\frac{(2-\beta^2)}{\beta}(1-\beta^2)^{\frac{1}{2}}-\frac{\pi}{2}\right]} \tag{69}$$

$$F_8(\beta)=\frac{(4/3)(1-\beta^2)^{\frac{1}{2}}}{\left[\tan^{-1}\frac{\beta}{(1-\beta^2)^{\frac{1}{2}}}+\frac{(2-\beta^2)}{\beta}(1-\beta^2)^{\frac{1}{2}}-\frac{\pi}{2}\right]}. \tag{70}$$

The four functions of β given are plotted in Fig. 3.

From these results we can understand the different types of information to be found from the various sorts of perturbations. We see from Figs. 1 and 2 that our expectation regarding needles is borne out by the calculation. A very thin needle corresponds to a very small value of β. In this case, F_1 is much greater than any of the other functions. That is, a large perturbation of frequency is produced by a needle pointing along the electric lines of force, but a very small perturbation if it is at right angles, and a very small magnetic perturbation no matter what its orientation. This means that a needle is a very useful device for finding the direction of the electric field at any point of the cavity. We shall see in the next section, however, that, since it is difficult to make a needle-shaped ellipsoid which is shaped accurately enough so that its quantitative measure of field is very useful, the major use of the needle seems to be that of determining direction.

From Fig. 3 we see that a thin disk (that is, with small β) has a negligible perturbing effect when it is normal to an electric field (F_6) or parallel to a magnetic field (F_8), whereas it has a large perturbing effect if it is parallel to an electric field (F_5) or normal to a magnetic field (F_7), as we expected. This allows us to find in a simple way the direction of the component of magnetic field normal to the electric field. First we establish the direction of the electric field, by using a needle, as in the preceding paragraph. Then we place a disk so that the direction of the electric field is in the plane of the disk, and rotate the disk about the direction of the electric field. The magnetic perturbation will be a maximum when the normal to the disk points along the component of magnetic field normal to the electric field. Furthermore, from the difference in perturbation in this orientation and the orientation at right angles, where the magnetic field is parallel to the plane of the disk, we can get the magnitude of the normal component of magnetic field, provided the disk is an ellipsoid of accurate dimension; this, as we shall see from the next section, is not hard to achieve. We have not yet determined the component of magnetic field parallel to the electric field (this component will of course be zero in many of the simpler types of cavity oscillation, in which electric and magnetic fields are perpendicular to each other, but this situation is by no means always the case). We can do it, however, in principle, as follows. The disk, when placed parallel to both electric and magnetic fields, measures only the electrical field. A sphere, however, measures the combination $E_a^2 - 1/2H_a^2$. Thus, comparison of the perturbations produced by the disk and sphere will give the magnitude not only of E, but of H as well. Since we have already determined the component of H normal to E by rotating the disk, this allows us to find the component of H parallel to E by subtraction. This procedure, although theoretically possible, would make rather severe demands on the accuracy of measurement, which as indicated in the next section is not likely to be high, and it might not be practicable.

A second technique, that of measuring the field components directly, is more tedious to apply, but it overcomes the necessity of an elaborate method of suspending the perturbing volume. For the general cavity, there will be three components of each E and H, corresponding to the projections along the desired coordinates axes. To measure these components, six independent equations relating them must be solved simultaneously at each point. The most convenient perturbing volumes would be a needle and a disk of very low β, each oriented successively along each coordinate axis. From the results of Secs. 2 and 3, it can be seen that these shapes and orientations lead to approximations which aid in the solution of the six simultaneous equations. Moreover, the problem of suspensions is greatly simplified since only the position of a given orientation is changed over a series of measurements. This measurement technique was used for the measurements described in the previous section and was found to be very effective.

4. EXPERIMENTAL VERIFICATION

The ease with which small changes in resonant frequency can be measured in a high-Q cavity resonator is one of the major justifications for developing this perturbation method of measuring field strengths. The problem is to measure a change in frequency of from a tenth to a hundredth of a megacycle in a resonant frequency of three thousand megacycles. The method employed here utilizes the stabilized microwave oscillator developed by R. V. Pound.[9] This circuit stabilizes a McNally klystron to a resonant cavity by means of a Magic-Tee discriminator feedback loop to the klystron reflector plate. The klystron is thus electronically tuned to the resonant frequency of the cavity and, when the circuit is properly designed, it can follow considerable shifts in the cavity frequency. The degree of stabilization obtained depends upon the Q of the cavity as well as the gain and stability of the feedback loops. To measure the frequency shifts produced by the perturbing volume, two stabilized oscillators were used, and a beat signal was obtained in the low-frequency band. One oscillator was stabilized with the cavity to be measured and the other with a high-Q tunable wavemeter, which gave control over the beat frequency obtained with no perturbation. This beat frequency was fed into an ordinary AM radio receiver, and by introducing the perturbing volume into the cavity whose fields were to be measured, the changes in resonant frequency were measured with the receiver. A block diagram of the system is given in Fig. 4. Since the stabilizing cavities and circuits were reasonably well insulated, both thermally and mechanically, the beat signal obtained was found to vary less than a kilocycle over a five- to ten-minute period. By using a well calibrated receiver with an expanded frequency dial, it was relatively simple to

[9] R. V. Pound, Rev. Sci. Instr. **17**, p. 490 ff. (1946).

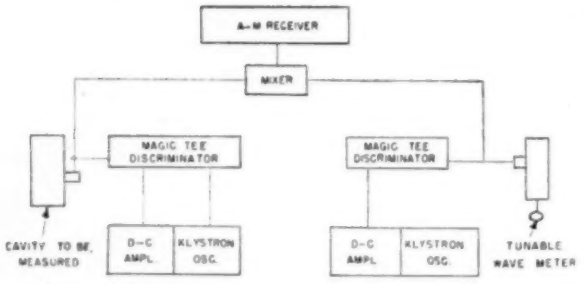


FIG. 4.

measure frequency shifts of one kilocycle in a beat frequency of approximately five megacycles. This permitted the use of quite small (compared to a wavelength) perturbing volumes, and eliminated much of the difficulty of the fields' varying rapidly over the perturbation.

For checking the theory previously developed, a cylindrical cavity of length L and radius R_0, operating in the TM_{010} mode, was used as one of the stabilizing cavities, and the frequency shifts measured for various shapes and sizes of perturbations. For this cavity,

$$\begin{aligned} L &= 7.620 \text{ cm} \\ R_0 &= 4.094 \text{ cm} \\ f_0 &= 2798.4 \text{ Mc} \end{aligned} \tag{71}$$

and

$$\begin{aligned} E_a^2 &= \frac{J_0^2(kr)}{\pi L R_0^2 J_1^2(kR_0)} \\ H_a^2 &= \frac{J_1^2(kr)}{\pi L R_0^2 J_1^2(kR_0)}. \end{aligned} \tag{72}$$

For the following measurements, the perturbing volume was cemented to a thin silk thread to suspend it in the cavity. The frequency shift of the thread could not be measured, which indicates the very slight effect it had on the fields. For a sphere in this cylindrical cavity,

$$E_a^2 - \tfrac{1}{2}H_a^2 = \frac{\omega_0^2 - \omega^2}{4\pi r_0^2 \omega_0^2} = \frac{2\Delta f}{4\pi r_0^2 f_0}. \tag{73}$$

Table I-A gives the results for three different spheres: First, in the E field alone on the axis of the cavity, and then at the maximum of $(E_a^2 - \frac{1}{2}H_a^2)$, with the percentage error of each measurement. From the low percentage error in the electric field measurements, it is seen that the perturbation theory checks to within a few percent, both the magnitude of the effect, and the cubic variation with sphere radius. For the measurements at the maximum of $(E_a^2 - \frac{1}{2}H_a^2)$, the percentage error is somewhat greater and can be accounted for by the difficulty in getting the sphere directly at the maximum and by the perturbation of the fields caused by the hole needed to introduce the spheres.

TABLE I-A spheres.

a (cm)	Δf (Mc)	E_a^2(cm^{-3}) (meas.) $\times 10^{-3}$	E_a^2(cm^{-3}) (calc.) $\times 10^{-3}$	Percentage of error	Δf (Mc)	$(E_a^2 - \frac{1}{2}H_a^2)$ (cm^{-3}) (meas.) $\times 10^{-3}$	$(E_a^2 - \frac{1}{2}H_a^2)$ (cm^{-3}) (calc.) $\times 10^{-3}$	Percentage of error
0.0493	5.09	9.10	9.20	1.3	0.693	1.24	1.31	5.1
0.0369	2.14	9.02	9.20	2.0	0.298	1.25	1.31	3.3
0.0246	0.645	9.15	9.20	0.6	0.090	1.28	1.31	1.9

Table I-B gives the results for the measurements made on disks. Four disks were used, three of very thin metal to check the variation with semi-major axis a and one of $\beta = 0.455$ to give some idea of the correlation of theory with variations of β. Because of the difficulty of aligning the disk in a mixed electric and magnetic field, and because disks are not as useful in measuring fields, the checking was done only for the electric field along the axis. The expression for the perturbation produced by a disk with its semi-major axis directed along the electric field is

$$\frac{\omega_0^2 - \omega^2}{\omega_0^2} = \frac{4\pi a^3 F_5(\beta)}{\omega_0^2} E_a^2 \tag{74}$$

TABLE I-B disks.

a (cm)	β	Δf (Mc)	$F_5(\beta)$	E_a^2(cm^{-3}) (meas.) $\times 10^{-3}$	E_a^2(cm^{-3}) (calc.) $\times 10^{-3}$	Percentage of error
0.0480	0.00	1.91	0.424	8.72	9.20	5.3
0.0370	0.00	0.92	0.424	9.10	9.20	1.3
0.0260	0.00	0.330	0.424	9.40	9.20	2.0
0.0477	0.48	2.52	0.690	7.20	9.20	22.6

from which

$$E_a^2 = \frac{2\Delta f}{4\pi a^3 F_5(\beta) f_0}. \tag{75}$$

The percentage errors for the thin disks are only slightly larger than for spheres and can easily be accounted for in the difficulty of correctly measuring a. The odd sizes of these thin disks result from the fact that they were punched from thin stock and the clearance between the inner and outer surfaces of the punch left burrs on the disks. The results of the $\beta = 0.38$ disk, however, have a large percentage error, and the reason for this is probably the very poor tolerances held in machining. As will be seen in the case of needles, the perturbation theory seems to be quite critical as to the cross section of the disks and needles. This $\beta = 0.48$ disk was examined under a microscope and its cross section clearly was not elliptic.

Table I-C gives the results for needles placed in the E_a field at the center of the cavity and at the value of the maximum of $F_2(\beta)E_a^2 - \frac{1}{2}F_4(\beta)H_a^2$. The needles were first aligned with their semi-major axis along the axis of the cavity and hence along the electric field.

TABLE I-C needles.

a (cm)	β	Δf (Mc)	$F_1(\beta)$	$E_a^2(\text{cm}^{-3})$ (meas.) $\times 10^{-3}$	(calc.) $\times 10^{-3}$	Percentage of error
0.0835	0.052	4.30	0.125	12.6	9.20	29.2
0.0755	0.101	3.54	0.165	10.7	9.20	16.9
0.0535	0.46	2.28	0.438	7.20	9.20	22.6
			$F_3(\beta)$			
0.0835	0.052	0.059	0.00182	11.8	9.20	28.4
0.0535	0.46	1.19	0.168	9.75	9.20	6.0

					$(F_3E_a^2-\frac{1}{2}F_4H_a^2)$ (cm^{-3})		
				$F_4(\beta)$	(meas.) $\times 10^{-3}$	(calc.) $\times 10^{-3}$	Percentage of error
0.0835	0.052	0.0125	0.00182	0.00358	0.0299	0.0294	1.4
0.0535	0.46	0.241	0.168	0.243	0.219	0.224	2.5

In this case

$$E_a^2=\frac{2\Delta f}{4\pi a^3F_1(\beta)f_0}. \tag{76}$$

In the second experiment, the needles were oriented with the semi-major axis perpendicular to the axis of the cavity, but still in the center of the cavity, and in this case

$$E_a^2=\frac{2\Delta f}{4\pi a^3F_3(\beta)f_0}. \tag{77}$$

In the third case, they were still oriented as in the second, but were displaced from the center to the maximum negative values of Δf. Then

$$F_3(\beta)E_a^2-\tfrac{1}{2}F_4(\beta)H_a^2=\frac{2\Delta f}{4\pi a^3f_0}. \tag{78}$$

From the percentage errors plotted in Table I-C, it is seen that there seems to be rather poor agreement with the theory, but, as in the case of disks, the reason is that the needles were not truly ellipsoids of revolution. Examination under a microscope showed that the $\beta=0.101$ needle was the closest to an ellipsoid, the $\beta=0.46$ next, and the $\beta=0.052$ the farthest, which agrees with the changes in percentage error. The experimental error also shows that the effects of the nonelliptic cross section are least important when the fields are at right angles to the a-axis and most important when they are parallel. Because of the machining difficulties, it was decided that it would not be worthwhile to attempt making needles of better cross section, although it should not be impossible to do so.

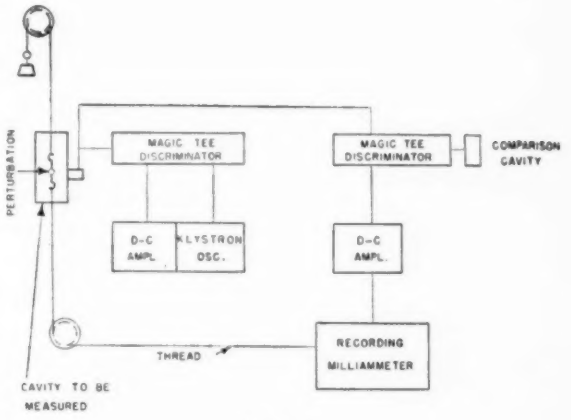


FIG. 5.

In the event that only the relative magnitude of field variations is of interest, there is a simple method of utilizing the perturbation measurements without the more laborious technique of finding the absolute field strengths. The schematic diagram of Fig. 5 represents the circuit used. The cavity whose fields are to be measured is used to stabilize a klystron oscillator whose output is fed into a Magic-Tee discriminator, a dc amplifier and to a milliammeter. If convenient, a recording milliammeter can be used, and the means of suspending the perturbing volume can be attached to a recording roll such that the recording milliammeter plots the relative field as a direct function of position. The choice of the proper perturbing shape and orientation will be dictated by the desired component of the field being measured. This system has also been used at some length, particularly where only a single field component is present, and has given very reproducible results.

Determination of Field Strength in a Linear Accelerator Cavity*

L. C. MAIER, JR.† AND J. C. SLATER
Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts
(Received August 2, 1951)

One theoretical and two experimental methods are described for finding the accelerating field in the M.I.T. linear accelerator cavity, in terms of the input power. One of the experimental methods is based on measuring the power leaking out through a calibrated hole in an end wall closing the cavity, the hole being located where the field strength is to be determined. The other method is based on the perturbation of the resonant frequency of the cavity by a small conducting sphere located on the axis. All three methods of determination check satisfactorily. In terms of the resulting relations, we discuss the probable field actually existing in the accelerator under operating conditions.

IN an earlier reference[1] it was shown that the accelerating field E (in volts per meter), in a linear accelerator, can be written in terms of P_0, the power per unit length fed into the accelerator (in watts per meter), of Q_0, the unloaded Q of the accelerator cavity, and of the wavelength λ_0 of operation (in meters), by the relation,

$$E=\alpha(377P_0Q_0/\lambda_0)^{\frac{1}{2}} \tag{1}$$

where α is a numerical factor depending only on geometry. It was stated (see reference 1, p. 491) that the calculated value of this quantity for the M.I.T. accelerator was 0.48. It is the purpose of the present paper to describe two independent experimental methods by which this calculated value of α was checked, and to discuss the accelerating field actually obtained in the accelerator by means of Eq. (1).

As a preliminary, we described the method used to calculate α. In reference 1, p. 486, we state that the electric field along the axis has the component $E\exp(j\omega(t-z/v))$ associated with the resonant wave, and stored energy per unit length $W=A\epsilon_0E^2\lambda_0^2$, where

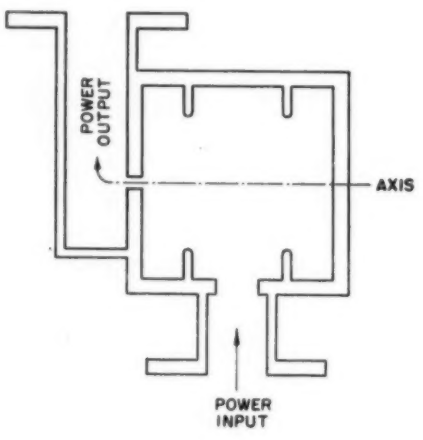


FIG. 1.

A is a numerical constant. Of course, the electric field is actually the real part of the exponential, or

$$E_0\cos\omega(t-z/v)=E_0(\cos\omega t\cos\omega z/v+\sin\omega t\sin\omega z/v),$$

where we have put a subscript 0 on E_0 to denote the value along the axis of the accelerator. At the same time, since we are dealing with a standing wave accelerator, there must be a traveling wave in the opposite direction. Superposing these two traveling waves, we have the standing wave $2E_0\cos\omega t\cos\omega z/v$ as the field along the axis. In addition to this standing wave, we shall have further components of shorter wavelength, which, however, are of negligibly small amplitude.

Now let us assume that the electric field at any point of the cavity is $E\cos\omega t$, where $E(x, y, z)$ is a function of position, which reduces to $2E_0\cos\omega z/v$ along the axis. Let $\langle E^2\rangle_{Av}$ be the average of E^2 throughout the cavity. Let us find the stored energy. The electric energy density is $\frac{1}{2}\epsilon_0E^2\cos^2\omega t$, and its time average is $\frac{1}{4}\epsilon_0E^2$. The average (over space) of the electric energy density is then $\frac{1}{4}\epsilon_0\langle E^2\rangle_{Av}$, so that the stored energy in unit length of the accelerator is this quantity times the volume per unit length. If we neglect the thickness of the irises, this volume is πR^2, where R is the radius of the accelerator. Thus the stored electric energy is $\frac{1}{4}\epsilon_0\langle E^2\rangle_{Av}\pi R^2$. Since the average magnetic energy equals the average electric energy, the total stored energy W per unit length is $\frac{1}{2}\epsilon_0\langle E^2\rangle_{Av}\pi R^2$.

We have already stated that $W=A\epsilon_0E_0^2\lambda_0^2$. Thus we see that $\frac{1}{2}\epsilon_0\langle E^2\rangle_{Av}\pi R^2=A\epsilon_0E_0^2\lambda_0^2$. Furthermore, we have $\alpha=(2\pi A)^{-\frac{1}{2}}$, by definition. Thus we find at once that

$$\alpha=(\lambda_0/\pi R)(E_0/\langle E^2\rangle_{Av}^{\frac{1}{2}}). \tag{2}$$

In other words, to find α, we must investigate the field inside the cavity, and find the ratio between the field along the axis, and the root mean square field throughout the cavity. In an earlier paper[2] the field inside a cavity of the type considered was calculated by a rather accurate approximation method. On the basis of that method, the ratio $E_0/\langle E^2\rangle_{Av}^{\frac{1}{2}}$ was evaluated numerically, and in this way the value of α already

* This work has been supported in part by the Signal Corps, the Air Materiel Command, and ONR.
† Now at Electronics Laboratory, General Electric Company, Syracuse, N. Y.
[1] J. C. Slater, Revs. Modern Phys. **20**, 473 (1948); see particularly p. 486 ff.
[2] J. C. Slater, "Electromagnetic waves in iris-loaded waveguides," Technical Report No. 48, Research Laboratory of Electronics, M.I.T. (1947).

quoted was found. We shall now proceed to describe the two experimental methods which have been used to check this theoretical value.

1. POWER MEASUREMENT METHOD FOR α

In this method, we used a section of cavity one wavelength long, as shown in Fig. 1, terminated by conducting planes at both ends. In the center of one end was a small hole, communicating with a wave guide output. The power coming out this hole was measured, and compared with the power fed into the cavity. The hole was so small as hardly to load the cavity at all. Thus we can assume that the power fed in is all dissipated in the walls. In this case the power coming out of the hole will be proportional to the square of the normal electric field strength at the hole, which is proportional to the quantity E_0 already used. The power fed in equals the power dissipated in the wall, which is ω/Q_0 times the stored energy. Thus if we know the properties of the output hole (the amount of power emerging, in its relation to the normal electric field at the hole), we have the material for calculating the ratio of E_0^2 to stored energy, which as we have seen is what we need for finding α. The calibration of the hole was carried out by making a similar experiment on a simple cylindrical cavity, as shown in Fig. 2, in which we can make an exact calculation of the field strength at the hole, and thus use a similar experiment to determine the properties of the hole. (Identical holes were used in both cases.)

First, let us find the field strength at the output hole for the accelerator. It is a point of maximum field along the z axis, so that in the earlier language the field will be $2E_0\cos\omega t$. If we assume that the power flowing out of the hole is a constant C, depending on the hole size, times the square of the peak field at the hole, this power out will be $4CE_0^2$. The value of E_0 is given by $\alpha(377PQ_0/L\lambda_0)^{\frac{1}{2}}$, where P is the power fed in, L the length of the section of accelerator, so that P/L is the power per unit length, which was called P_0 in Eq. (1). Thus the ratio of power out to power in is

$$\text{Power out/power in} = 4C\alpha^2(377Q_0/L\lambda_0), \qquad (3)$$

where Q_0 and L refer to the section of accelerator.

Now, in a similar way let us find the ratio of power out to power in for the cylindrical cavity. For this cavity the electric field is independent of z, points along the z axis, and is proportional to $J_0(kr)$, where $k=2\pi/\lambda_0$. This field goes to zero when $kr=2.405$, the first zero of J_0, which occurs at the radius R of the cavity. Thus if the electric field at the hole is $E_0\cos\omega t$, the field at any interior point is $E_0\cos\omega t J_0(kr)$, since $J_0(0)=1$. The stored energy, by the type of argument used before, is $\frac{1}{2}\epsilon_0E_0^2\int J_0^2(kr)dv$, the integration being carried over the volume of the cavity. This integral is

$$L\int_0^R 2\pi r J_0^2(kr)dr = (2\pi L/k^2)\int_0^{2.405} zJ_0^2(z)dz.$$

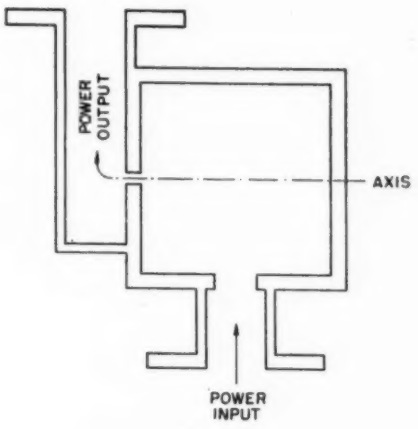


FIG. 2.

However, we have the formula,

$$\int zJ_0^2(z)dz = (z^2/2)[J_0^2(z)+J_1^2(z)].$$

Since $J_0(2.405)=0$, this gives $(2.405)^2J_1^2(2.405)/2$. We have $J_1(2.405)=0.519$. Hence, the stored energy is $\frac{1}{2}\epsilon_0E_0^2(2\pi L/k^2)(2.405)^2(0.519)^2/2$. We can use the relation $2\pi R/\lambda_0=2.405$ to reduce this to the value of $\frac{1}{2}\epsilon_0E_0^2(0.519)^2\pi R^2L$. This stored energy equals the ratio PQ_0/ω, where P is the input power, Q_0 the unloaded Q, of the cavity. In this case the power output will be CE_0^2, since the peak field at the hole is E_0. Putting these pieces of information together, we have

$$\text{Power out/power in} = CE_0^2/P = 2CQ_0/\omega\epsilon_0(0.519)^2\pi R^2L$$
$$= (CQ_0/L)(\lambda_0/2\pi c)(2/\epsilon_0(0.519)^2\pi R^2).$$

We use the relations $c=(\epsilon_0\mu_0)^{-\frac{1}{2}}$, 377 ohms $=(\mu_0/\epsilon_0)^{\frac{1}{2}}$, and find

$$\text{Power out/power in} = (377CQ_0/L)\lambda_0/((0.519)^2\pi^2R^2).$$

We use $2\pi R/\lambda_0=2.405$, and find

$$\text{Power out/power in} = (377CQ_0/L\lambda_0)(2/(2.405)(0.519))^2$$
$$= 2.56(377CQ_0/L\lambda_0) \qquad (4)$$

where this Q_0, L refer to the cylindrical cavity.

We have now measured the ratio of power out to power in for both types of cavity. We can then combine the two equations, and have

$$\alpha^2 = (2.56/4)[(L/Q_0)(\text{power out/power in})]_{\text{acc}}$$
$$\div[(L/Q_0)(\text{power out/power in})]_{\text{cyl}}.$$

The experimental values are as follows. For the section of accelerator, the length is one wavelength, Q_0 is 13,825, and the power out is 41.4 db below the power in; that is, power out/power in $=10^{-4.14}$. For the cylinder, the length is 2.074 inches $=5.26$ cm, Q_0 is 8390, and the power out is 36.54 db below the power in, or power

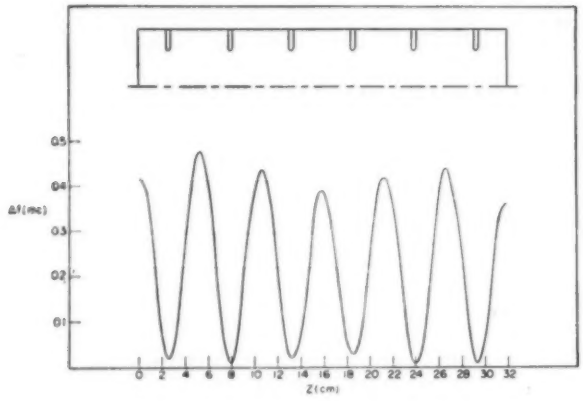


FIG. 3.

out/power in $= 10^{-3.654}$. Thus we have

$$\alpha^2 = (2.56/4)(10.69/13{,}825)10^{-4.142} \div [(5.26/8390)10^{-3.654}]$$

$$\alpha = 0.51$$

in satisfactory agreement with the calculated value 0.48.

2. PERTURBATION METHOD

In another reference[3] it is shown that the insertion of a small conducting ellipsoid of revolution into a resonant electromagnetic cavity will perturb the resonant frequency by an amount depending upon the shape and size of the ellipsoid and the local values of the electric and magnetic fields. Relations are developed of the form,

$$(\omega^2 - \omega_0^2)/\omega_0^2 = (4\pi/3)a^3[3F_E(\beta)E_a^2 - \tfrac{3}{2}F_H(\beta)H_a^2] \quad (5)$$

where E_a and H_a are the electric and magnetic fields at the point of insertion, but normalized so that the integral over the volume is unity; ω_0 is the unperturbed resonant frequency; ω is the perturbed resonant frequency; $F_E(\beta)$ and $F_H(\beta)$ are functions of the shape and orientation of the perturbing ellipsoids; and a is its semi-major axis. Experimental measuring techniques are also given. This method has been used to explore the electric and magnetic in a typical section[4] of the M.I.T. accelerator.

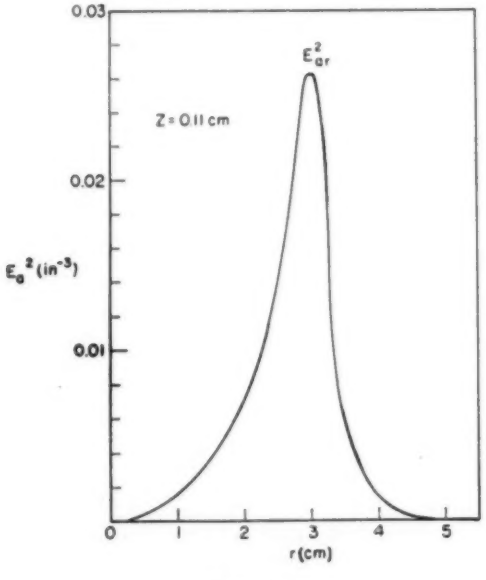


FIG. 4.

Two types of measurements were made. The simpler method was to explore the field along the axis by means of a spherical object. Along the axis the field must be only electric, by symmetry, it must be longitudinal, and it should vary sinusoidally with distance, so that a spherical object can verify the sinusoidal dependence, and give the required maximum amplitude. Since this is all that is needed to describe the actual operation of the accelerator, in which the electrons travel along the axis, this type of measurement is the most important one for practical purposes. A typical experimental curve for frequency shift Δf as a function of distance along the axis is given in Fig. 3, for a section six half-wavelengths long, with the positions of the irises and end walls plotted to the same scale. We see that the sinusoidal shape is perturbed near the middle; this is a result of the fact that the section had a rather large hole in one side, at this position, for feeding power in. From the maximum amount of the frequency shift in Fig. 3, we can find the field strength at the axis, which is what is needed for calculating α and hence the acceleration of the electron.

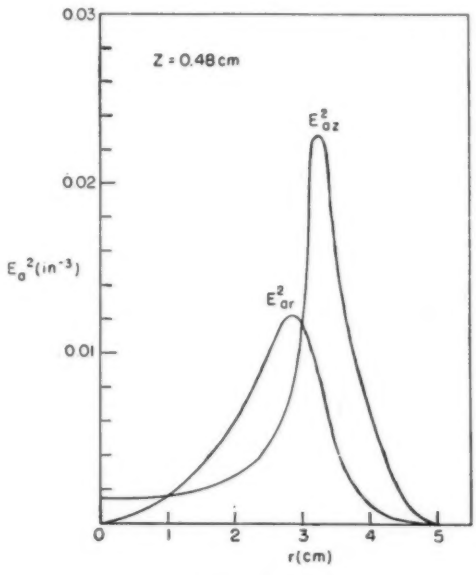


FIG. 5.

Measurements like those of Fig. 3 were made as a matter of routine test, on all the sections constructed for

[3] L. C. Maier, Jr., and J. C. Slater, J. Appl. Phys. **23**, 68 (1952).

[4] Demos, Kip, and Slater, "The M.I.T. linear accelerator," Technical Report No. 203, Research Laboratory of Electronics, M.I.T. (unpublished).

the accelerator.[4] These sections were eight wavelengths long, and were teminated at the ends by drift tubes, with holes large enough to pass the electron beam, although too small for the field to penetrate. The drift tubes introduced considerable perturbations into the end sections, which were compensated as well as possible by choosing the dimensions properly in manufacture. However, errors in dimensions sometimes crept in which resulted in a lack of symmetry between the two ends of the section, and in such cases the excitation sometimes would be much stronger at one end than at the other, as was immediately revealed by curves similar to Fig. 3. Modifications were then made to the dimensions of the drift tubes, to correct this asymmetry, and the "ball test" resulting from running the sphere through the cavity and measuring the frequency change proved to be invaluable in attaining uniform excitation.

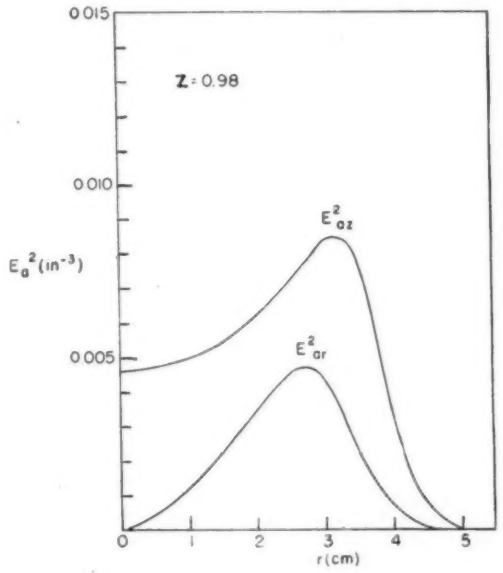


FIG. 6.

The test could be used in other ways as well. Near the π-mode in which this accelerator is intended to operate, there are other modes of neighboring frequency but quite different wavelength. The nature of these modes was established beyond doubt by exciting them, and running the ball down them in the other modes; the wavelength could be immediately determined. These other modes included not only other members of the TM series of modes, but also a series of TE modes, whose frequencies lay in the same region. These also were identified by the ball test, and their wavelengths found, from which a relation between wavelength and frequency was set up empirically for them. As a result of this relation, it was possible to choose a length for the accelerating section (eight wavelengths of the π-mode) for which no TE modes lay in the immediate neighborhood of the π-mode, although other lengths almost all happened to have TE modes uncomfortably close. The

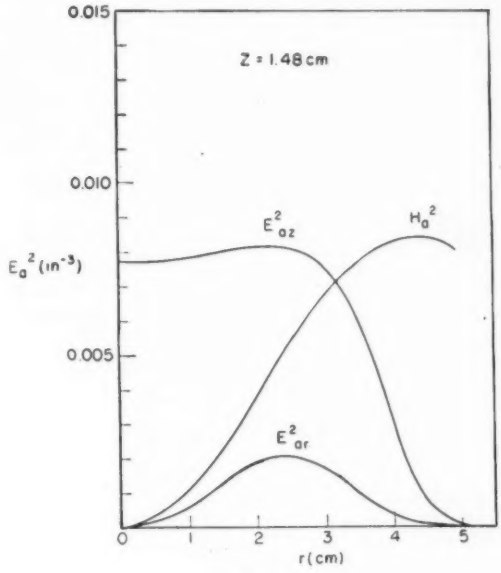


FIG. 7.

TE modes were identified as such by inserting a needle rather than a ball: the maximum perturbation of frequency came when the needle was transverse to the axis, rather than along the axis as in the TM modes.

The other type of measurement made on the linear accelerator cavity was a more elaborate one, designed to give all three components of field, E_r, E_z, E_θ, at all points of the cavity, in the π-mode. The details of the method are given in reference 5. Briefly, measurements were made with a sphere, and with a needle in two orientations, along the radius r and along the axis z,

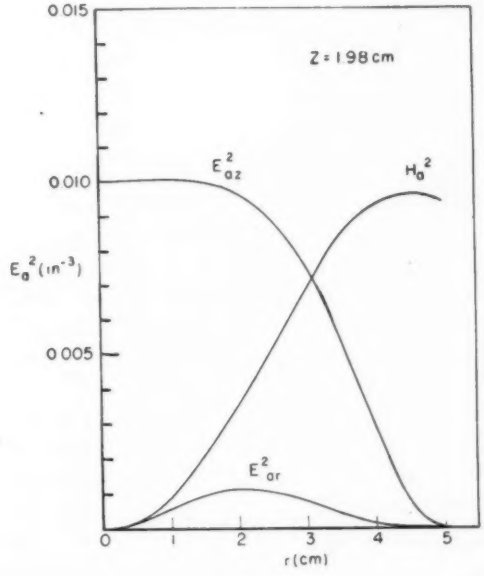


FIG. 8.

[5] L. C. Maier, Jr., "Field strength measurements in resonant cavities," Technical Report No. 143, Research Laboratory of Electronics, M.I.T. November 2, 1949.

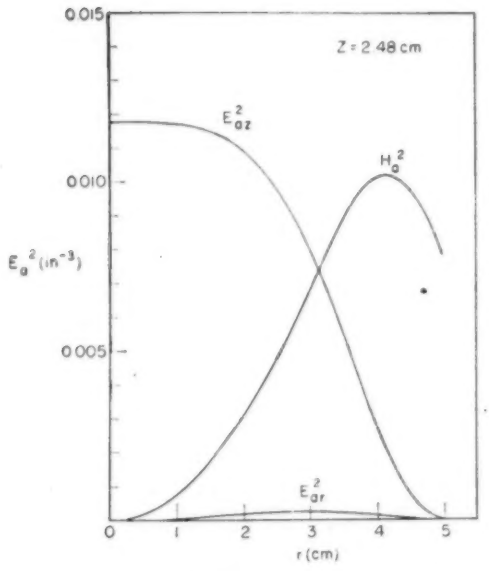


FIG. 9.

at a large number of points within the cavity. If the needle were thin enough, then as we saw in Sec. 2 (reference 3) we should get E_r and E_z from its two orientations, and then should get H_θ by combining these measurements with those from the sphere. Since the needle is not infinitely thin, the method is somewhat more complicated, in that each measurement gives a result depending on all three components, but the situation is not greatly different from the simple one just described.

From these measurements, the results of Figs. 4 to 10 were computed. Each of these gives the value of E_{ar}^2, E_{az}^2, $H_{a\theta}^2$, as a function of r, the distance out from the axis of the accelerator (the wall corresponds to $r=5.08$ cm, the radius of the hole in the iris to 3.18 cm). Each one is computed for a different value of z, the distance along the accelerator, measured from one of the irises; the midpoint between irises corresponds to $z=2.68$ cm, in Fig. 10. We see that the curves have the right general behavior. Close to the iris (Fig. 4), E_r rises to a very high value as we get near the radius of the hole in the iris, since the field is extremely large just at the iris itself; it then falls to a very small value as r gets larger than the hole in the iris, since the radial component E_r is then tangential to the surface of the iris, and must vanish just at the surface. It was not practicable to measure E_z, and hence H, so close to the iris. In Fig. 5 some distance further from the iris, E_r has diminished. Here a measurement of E_z could be made, indicating the very strong component at about $r=3$ cm pointing into the edge of the iris, much stronger than along the axis, for $r=0$. As z increases, going from Fig. 6 to 10, E_r rapidly decreases, being of course zero at the midpoint between irises, and E_z correspondingly increases. We see that as we get away from the immediate influence of the iris, E_z stays remarkably constant for a considerable range of r, almost out to the radius of the iris. This is to be expected theoretically: it is easily shown that a TM wave of this type, traveling faster than the velocity of light, will have its E_z decrease quadratically with r, as we go away from the axis (being given by a Bessel function), while, if it is going slower than the velocity of light, E_z will increase quadratically; for just the velocity of light, the first term in the dependence of E_z on r will be a fourth-power term. The value of E_z for $r=0$ of course is the quantity mentioned in the first paragraph of this section, and illustrated by Fig. 3.

In addition to the electric field, we give H_a in Figs. 4 to 10. It was not practicable to calculate it for smaller values of z than that of Fig. 7; since it is found by subtraction, the values found for it will be very inaccurate in case the magnetic field is much less than the electric field (both being assumed normalized), as we have near the iris. This is a drawback of the perturbation method, in which there is no direct measurement of the magnetic field alone. The general form of the magnetic field is as we should expect. By Ampere's law we know that the line integral of H around a circle of constant r is proportional to the surface integral of E_z (proportional to the normal component of displacement current) over the circle. That is, for small r, where E_z is approximately constant, rH is proportional to r^2, or H is proportional to r, H^2 to r^2, as the figures show. As we approach the maximum radius of the cavity, the surface integral becomes constant, since E_z drops to zero, so that rH approaches constancy, which means that H itself must go to a maximum and fall again, as it is seen to do.

We shall now show how these measurements can be used to find α. The frequency perturbation depends on the value of E_a on the axis, and on account of the normalization of E_a, this determines the ratio of E on the

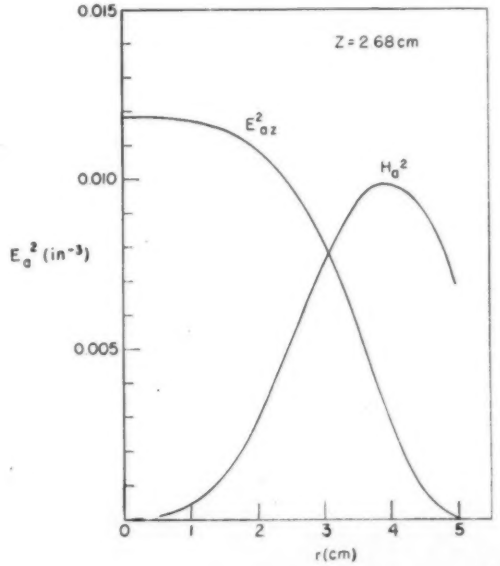


FIG. 10.

axis to the root mean square E throughout the cavity, which is what is needed for determining α. This measurement then gives an absolute measurement of α.

Let the length of accelerator cavity used for the ball test be L, its radius R. Let $\langle E_a^2\rangle_{Av}$ be the mean square of E_a throughout the cavity. On account of the normalization, we then have $\langle E_a^2\rangle_{Av}=1/(\pi R^2 L)$. We know that the maximum value of the electric field along the axis is $2E_0$, which corresponds to the maximum of E_a which we must use in the forgoing formula for finding the maximum frequency displacement. On account of the proportionality between E and E_a, we have

$$2E_0/E_a=(\langle E^2\rangle_{Av}/\langle E_a^2\rangle_{Av})^{\frac{1}{2}}.$$

Thus the value of E_a^2 on the axis is $(4E_0^2/\langle E^2\rangle_{Av})/(\pi R^2 L)$. We can now use our earlier formula for α to rewrite this. We have Eq. (5), and using it for a sphere ($F_E(\beta)=1.0$) in regions where only an electric field exists, we find

$$\omega^2=\omega_0^2(1-(16\pi^2 r^3\alpha^2/L\lambda_0^2)).$$

From this, if the frequency is f_0, change in frequency Δf, we have

$$\alpha=-(\lambda_0^3\Delta f L/8\pi^2 r^3 c)^{\frac{1}{2}}$$

where c is the velocity of light. The measurements were made with a ball of diameter $\frac{1}{4}$ inch, or radius 0.125 inch$=0.317$ cm. The length of section of accelerator used was three wavelengths, or 3×10.675 cm, for the section used. The maximum frequency shift from Fig. 3 is 0.434 Mc. Inserting these values, we find $\alpha=0.474$, in very good agreement with the calculated value of 0.48. From the agreement of this result, and the other experimental measurement of Sec. 1, we may conclude that the theoretical value 0.48 well represents the facts.

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 1 JANUARY, 1952

Forced Thickness-Shear and Flexural Vibrations of Piezoelectric Crystal Plates

RAYMOND D. MINDLIN
Department of Civil Engineering, Columbia University, New York
(Received July 10, 1951)

An approximate theory is given for the vibrations of piezoelectric crystal plates in thickness-shear and flexural modes, including the interaction between the elastic and electric fields. The problem of the rectangular, AT, quartz plate is solved and computed frequencies are compared with measurements by Sykes. Algebraic formulas are obtained relating resonant frequencies with dimensions, elastic and electric constants, and orientation of cut.

I. INTRODUCTION

IN two previous papers procedures were described for studying various aspects of shear and flexural motions of elastic plates. In the first paper[1] (referred to in the sequel as I) the three-dimensional equations of isotropic elasticity were reduced to two-dimensional, plate-flexure equations and, in the course of the reduction, the terms representing rotatory inertia and shear were retained. It was shown that these terms account for the discrepancy between the velocities of short waves, predicted by the classical theory of plates and the complete three-dimensional theory. As a consequence, it was concluded that the improved plate theory must contain high frequency vibration phenomena beyond the range of applicability of classical plate theory. In the second paper[2] (referred to in the sequel as II) the improved plate theory was extended to include crystal plates. A problem of vibrations of a rectangular plate was solved and computations of frequencies were made for AT-cut quartz plates having finite length to thickness ratios. The theory revealed, both qualitatively and quantitatively, the intricate high frequency spectrum observed in the laboratory.

In the present paper the improved plate theory is further extended to accommodate the piezoelectric relations and the electric field equations. The theory includes, then, the interaction between the elastic and electric fields in a vibrating piezoelectric crystal.

The problem of the rectangular plate, analogous to that considered in II, is reconsidered as one of piezoelectrically forced vibrations. The electric-elastic equations are solved simultaneously and the boundary conditions yield a transcendental frequency equation of the same form as that found in II. The effect of the interaction of the two fields on the frequencies is contained in two correction terms: one on the thickness-shear modulus and one on the plate-flexure modulus. One of the results is a formula for the surface charge density, in terms of the shear strain and the surface potential, from which the current through the vibrating plate may be calculated.

The main advantage gained through the use of an adequate, two-dimensional theory of plates is that the mathematical solution for a plate of finite length is no more difficult than the solution for an infinite plate in

[1] R. D. Mindlin, J. Appl. Mech. **18**, No. 1, 31–38 (March, 1951).
[2] R. D. Mindlin, J. Appl. Phys. **22**, 316–323 (1951).

the three-dimensional theory. This is done at the sacrifice of higher modes through the thickness; but the crystalline nature of the medium, the length modes, and the coupling between thickness-shear and flexural modes are retained, and these appear to contain the essence of the observed phenomena.

THREE-DIMENSIONAL EQUATIONS

It is expedient to record, first, the three-dimensional equations of the elastic and electric fields and the piezoelectric relations connecting them. The elastic and electric properties of the material are assumed to have monoclinic, sphenoidal symmetry, with the x-axis an axis of binary symmetry. Then the piezoelectric relations are

$$\begin{aligned}
T_1&=c_{11}S_1+c_{12}S_2+c_{13}S_3+c_{14}S_4-e_{11}E_1,\\
T_2&=c_{12}S_1+c_{22}S_2+c_{23}S_3+c_{24}S_4-e_{12}E_1,\\
T_3&=c_{13}S_1+c_{23}S_2+c_{33}S_3+c_{34}S_4-e_{13}E_1,\\
T_4&=c_{14}S_1+c_{24}S_2+c_{34}S_3+c_{44}S_4-e_{14}E_1,\\
T_5&=c_{55}S_5+c_{56}S_6-e_{25}E_2-e_{35}E_3,\\
T_6&=c_{56}S_5+c_{66}S_6-e_{26}E_2-e_{36}E_3.
\end{aligned} \tag{1}$$

$$\begin{aligned}
P_1&=e_{11}S_1+e_{12}S_2+e_{13}S_3+e_{14}S_4+\kappa_{11}E_1,\\
P_2&=e_{25}S_5+e_{26}S_6+\kappa_{22}E_2+\kappa_{23}E_3,\\
P_3&=e_{35}S_5+e_{36}S_6+\kappa_{23}E_2+\kappa_{33}E_3,
\end{aligned} \tag{2}$$

where T_i, S_i, E_i, P_i are the components of stress, strain, electric force, and polarization, respectively, and c_{ij}, e_{ij}, κ_{ij} are the elastic, piezoelectric, and susceptibility coefficients, respectively.

The strain-displacement relations are

$$\begin{aligned}
S_1&=\partial u_1/\partial x, \quad S_4=\partial u_3/\partial y+\partial u_2/\partial z,\\
S_2&=\partial u_2/\partial y, \quad S_5=\partial u_1/\partial z+\partial u_3/\partial x,\\
S_3&=\partial u_3/\partial z, \quad S_6=\partial u_2/\partial x+\partial u_1/\partial y.
\end{aligned} \tag{3}$$

The stress equations of motion are

$$\begin{aligned}
\frac{\partial T_1}{\partial x}+\frac{\partial T_6}{\partial y}+\frac{\partial T_5}{\partial z}&=\rho\frac{\partial^2 u_1}{\partial t^2},\\
\frac{\partial T_6}{\partial x}+\frac{\partial T_2}{\partial y}+\frac{\partial T_4}{\partial z}&=\rho\frac{\partial^2 u_2}{\partial t^2},\\
\frac{\partial T_5}{\partial x}+\frac{\partial T_4}{\partial y}+\frac{\partial T_3}{\partial z}&=\rho\frac{\partial^2 u_3}{\partial t^2},
\end{aligned} \tag{4}$$

and the pertinent electric field equations are

$$(E_1, E_2, E_3)=-\left(\frac{\partial V}{\partial x}, \frac{\partial V}{\partial y}, \frac{\partial V}{\partial z}\right), \tag{5}$$

$$\frac{\partial^2 V}{\partial x^2}+\frac{\partial^2 V}{\partial y^2}+\frac{\partial^2 V}{\partial z^2}=4\pi\left(\frac{\partial P_1}{\partial x}+\frac{\partial P_2}{\partial y}+\frac{\partial P_3}{\partial z}\right), \tag{6}$$

where V is the electrostatic potential. The influence of the magnetic field is omitted.

Through the use of Eqs. (1), (2), (3), and (5), Eqs. (4) and (6) may be expressed in terms of the dependent variables u_1, u_2, u_3, and V. Thus, consideration of the electric field adds one differential equation and one dependent variable to the usual equations of elasticity.

DERIVATION OF PLATE EQUATIONS

As in II, the faces of the plate are the planes $y=\pm h/2$ of an x, y, z system of rectangular coordinates. The normal components of the tractions on planes parallel to the faces are neglected, i.e.,

$$T_2=0 \tag{7}$$

throughout the plate. In addition the faces are assumed to be traction-free, so that

$$T_6=T_4=0 \quad \text{on} \quad y=\pm h/2. \tag{8}$$

Five plate-stress components (three stress couples and two stress resultants) are defined, in the usual manner, by

$$\begin{aligned}
(M_x, M_z, M_{xz})&=\int_{-h/2}^{h/2}(T_1, T_3, T_5)y\,dy,\\
(Q_x, Q_z)&=\int_{-h/2}^{h/2}(T_6, T_4)\,dy.
\end{aligned} \tag{9}$$

In addition, for the purposes of this paper, three plate-polarization components are defined by

$$\begin{aligned}
(P_x, P_z)&=\int_{-h/2}^{h/2}y(P_1, P_3)\,dy,\\
P_y&=\int_{-h/2}^{h/2}P_2\,dy.
\end{aligned} \tag{10}$$

(In order to avoid a new symbol, true polarization components are designated by the letter P with number subscripts, while plate-polarization components are designated by the letter P with letter subscripts.)

The four dependent variables of the three-dimensional theory are now assumed to have the form

$$hu_1=y\xi, \quad u_2=\eta, \quad hu_3=y\zeta, \quad hV=y\varphi, \tag{11}$$

where ξ, η, ζ and φ are functions of x, z and t only.

The conversion of the three-dimensional piezoelectric relations (1) and (2) to two dimensions proceeds as follows. Setting $T_2=0$, the second of (1) is solved for S_2 and this is used to eliminate S_2 from the remaining five of (1) and from (2). In the resulting eight equations, the strain and electric force components are expressed in terms of the displacements and electric potential in accordance with (3) and (5). The latter, in turn, are expressed in terms of ξ, η, ζ, and φ through (11). Integrations are now performed on the eight equations so as to convert the left hand sides to plate-stress and plate-polarization components, in accordance with (9)

and (10), with the result:

$$\begin{aligned} M_x &= D_1\partial\xi/\partial x + D_2\partial\zeta/\partial z + F_1\partial\varphi/\partial x, \\ M_z &= D_2\partial\xi/\partial x + D_3\partial\zeta/\partial z + F_2\partial\varphi/\partial x, \\ M_{xz} &= D_5(\partial\xi/\partial z + \partial\zeta/\partial x) + F_5\partial\varphi/\partial z, \\ Q_z &= D_4(\zeta + h\partial\eta/\partial z), \\ Q_x &= D_6(\xi + h\partial\eta/\partial x) + F_6\varphi. \end{aligned} \tag{12}$$

$$\begin{aligned} P_x &= F_1\partial\xi/\partial x + F_2\partial\zeta/\partial z - C_1\partial\varphi/\partial x, \\ P_y &= F_6(\xi + h\partial\eta/\partial x) - C_2\varphi, \\ P_z &= F_5(\partial\xi/\partial z + \partial\zeta/\partial x) - C_3\partial\varphi/\partial z, \end{aligned} \tag{13}$$

where

$$\begin{aligned} D_1 &= (c_{11} - c_{12}^2/c_{22})R, & D_2 &= (c_{13} - c_{12}c_{23}/c_{22})R, \\ D_3 &= (c_{33} - c_{23}^2/c_{22})R, & D_4 &= \kappa^2(c_{44} - c_{24}^2/c_{22}), \\ D_5 &= c_{55}R, & D_6 &= \kappa^2 c_{66}, \\ F_1 &= (e_{11} - c_{12}e_{12}/c_{22})R, & F_2 &= (e_{13} - c_{23}e_{12}/c_{22})R, \\ F_5 &= e_{35}R, & F_6 &= e_{26}, \\ C_1 &= (\kappa_{11} + e_{12}^2/c_{22})R, & C_2 &= \kappa_{22}, \quad C_3 = \kappa_{33}R, \end{aligned} \tag{14}$$

and

$$R = h^2/12.$$

As in I and II, the factor κ^2 is introduced in the constants D_4 and D_6 in order to accommodate, as well as possible, the assumed simple y-dependence to the true y-dependence which varies markedly with wavelength and frequency. As before, the value of κ^2 will be determined by comparison of an exact solution of the three-dimensional equations with the corresponding solution of the plate equations.

The three-dimensional differential equations are converted to plate equations by integrating over the thickness of the plate after multiplying (6) and the first and third of (4) by y. Using Eqs. (7) to (11), the result is

$$\begin{aligned} \frac{\partial M_x}{\partial x} - Q_x + \frac{\partial M_{xz}}{\partial z} &= \rho R\frac{\partial^2\xi}{\partial t^2}, \\ \frac{\partial Q_x}{\partial x} + \frac{\partial Q_z}{\partial z} &= \rho h\frac{\partial^2\eta}{\partial t^2}, \\ \frac{\partial M_{xz}}{\partial x} - Q_z + \frac{\partial M_z}{\partial z} &= \rho R\frac{\partial^2\zeta}{\partial t^2}, \\ R\left(\frac{\partial^2\varphi}{\partial x^2} + \frac{\partial^2\varphi}{\partial z^2}\right) &= 4\pi\left(\frac{\partial P_x}{\partial x} + \frac{\partial P_z}{\partial z}\right). \end{aligned} \tag{15}$$

Finally, the plate-stress and plate-polarization components may be eliminated from (15), through the use of the plate-piezoelectric relations (12) and (13), leaving four differential equations governing the four dependent variables ξ, η, ζ, φ.

BOUNDARY CONDITIONS

If the electric terms are omitted in (12), (13), and (15), the equations reduce to those found in II. It was shown in I how the boundary conditions appropriate to the latter equations may be obtained from a plate-analogue of Neumann's uniqueness theorem. In terms of components referred to coordinates ν and s, measured, respectively, normal to and along the boundary of the plate, it was found that, at each point on the edge of the plate, there must be specified one member of each of the three products

$$M_\nu\frac{\partial\xi_\nu}{\partial t}, \quad M_{\nu s}\frac{\partial\xi_s}{\partial t}, \quad Q_\nu\frac{\partial\eta}{\partial t}, \tag{16}$$

where ξ_ν and ξ_s are the components of the vector (ξ, ζ) in the ν and s directions. On a free edge, for example,

$$M_\nu = M_{\nu s} = Q_\nu = 0. \tag{17}$$

To account for the electric terms it is only necessary to add the electrostatic energy, V_e, and proceed as before. Thus,

$$\begin{aligned} 8\pi V_e &= \int\int\int (E_1D_1 + E_2D_2 + E_3D_3)dxdydz \\ &= \int\int\int (E_1^2 + E_2^2 + E_3^2)dxdydz \\ &\quad + 4\pi\int\int\int (E_1P_1 + E_2P_2 + E_3P_3)dxdydz. \end{aligned} \tag{18}$$

From (5) and the last of (11), this becomes

$$\begin{aligned} 8\pi V_e &= h^{-2}\int\int\int\left[\left(y\frac{\partial\varphi}{\partial x}\right)^2 + \varphi^2 \right. \\ &\quad \left. + \left(y\frac{\partial\varphi}{\partial z}\right)^2\right]dxdydz - 4\pi h^{-1} \\ &\quad \times\int\int\int\left(P_1y\frac{\partial\varphi}{\partial x} + P_2\varphi + P_3y\frac{\partial\varphi}{\partial z}\right)dxdydz. \end{aligned} \tag{19}$$

Integrating over the thickness of the plate and using (10), we have

$$\begin{aligned} 8\pi V_e &= h^{-1}\int\int\left[R\left(\frac{\partial\varphi}{\partial x}\right)^2 + \varphi^2 + R\left(\frac{\partial\varphi}{\partial z}\right)^2\right]dxdz \\ &\quad - 4\pi h^{-1}\int\int\left(P_x\frac{\partial\varphi}{\partial x} + P_y\varphi + P_z\frac{\partial\varphi}{\partial z}\right)dxdz. \end{aligned} \tag{20}$$

The four terms containing derivatives are integrated by parts to obtain

$$\begin{aligned} 8\pi h V_e &= \int\int\left[4\pi\left(\frac{\partial P_x}{\partial x} + \frac{\partial P_z}{\partial z}\right)\right. \\ &\quad \left. - R\left(\frac{\partial^2\varphi}{\partial x^2} + \frac{\partial^2\varphi}{\partial z^2}\right)\right]\varphi dxdz \\ &\quad - \int\int(4\pi P_y - \varphi)\varphi dxdz \\ &\quad - \oint\left(4\pi P_\nu - R\frac{\partial\varphi}{\partial\nu}\right)\varphi ds. \end{aligned} \tag{21}$$

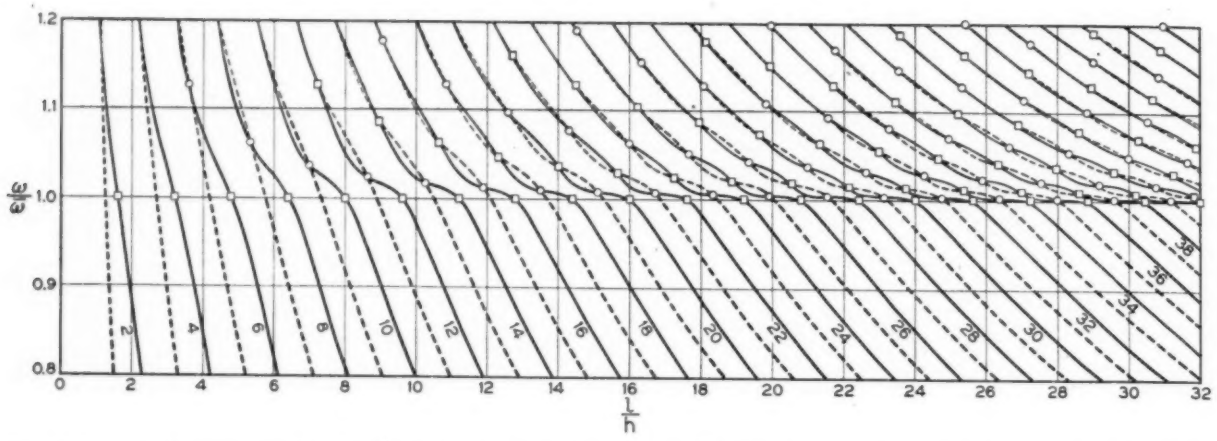


FIG. 1. Resonant frequencies of an AT-cut, rectangular, quartz plate as calculated from Eqs. (42), (43), and (44). For the circled and squared points, at which resonant frequencies of even and odd modes coincide, the transcendental frequency equation degenerates and the coordinates are given by simple algebraic formulas.

Noting that the first integral vanishes because of the last of (15) and writing

$$4\pi h\sigma_s = 4\pi P_y - \varphi, \qquad 4\pi h^2\sigma_e = 4\pi P_\nu - R\partial\varphi/\partial\nu, \tag{22}$$

we have

$$2V_e = -\int\int \sigma_s\varphi dxdz - h\oint \sigma_e\varphi ds. \tag{23}$$

It will be observed that σ_s is the surface charge density, σ_e is an "edge charge density," and φ is the potential difference between the two surfaces of the plate. Thus, by the argument used in I, the additional boundary conditions are:

1. At each point on the surface of the plate, either φ or σ_s may be specified.
2. At each point on the edge of the plate, either φ or σ_e may be specified.

A useful formula for the surface charge density may be obtained from (22), (13), and (14):

$$h\sigma_s = F_6(\xi + h\partial\eta/\partial x) - (1+4\pi\kappa_{22})\varphi/4\pi.$$

It is convenient to introduce the dielectric constants

$$K_i = (1+4\pi\kappa_{ii})/4\pi, \quad i=1,2 \tag{24}$$

so that

$$h\sigma_s = F_6(\xi + h\partial\eta/\partial x) - K_2\varphi. \tag{25}$$

INFINITE PLATE: DETERMINATION OF κ^2

The solution of the three-dimensional equations for the circular frequency, ω, of thickness-shear vibrations of an infinite plate with infinite plated electrodes is given by[3]

$$\omega^2 = \frac{4\alpha_n{}^2 c_{66}}{\rho h^2}\left(1+\frac{e_{26}{}^2}{c_{66}K_2}\right), \tag{26}$$

where α_n are the roots of

$$\tan\alpha_n = \alpha_n\left(1+\frac{c_{66}K_2}{e_{26}{}^2}\right). \tag{27}$$

The corresponding solution of the plate equations is obtained by setting

$$\varphi = \varphi_0\exp(i\tilde{\omega}t), \quad \xi = \xi_0\exp(i\tilde{\omega}t), \quad \eta = \zeta = 0,$$

where φ_0 and ξ_0 are constants. Then, from (12) and (13),

$$M_x = M_z = M_{xz} = Q_z = P_x = P_z = 0,$$
$$Q_x = (D_6\xi_0 + F_6\varphi_0)\exp(i\tilde{\omega}t),$$
$$P_y = (F_6\xi_0 - C_2\varphi_0)\exp(i\tilde{\omega}t).$$

Hence the last three of (15) are satisfied identically and the first gives

$$\xi_0 = F_6\varphi_0/(\rho\tilde{\omega}^2R - D_6),$$

which is the displacement as a function of impressed voltage and frequency and the elastic and piezoelectric constants. Resonance occurs when

$$\tilde{\omega}^2 = D_6/\rho R = 12\kappa^2 c_{66}/\rho h^2. \tag{28}$$

Comparing (26) and (28), we find

$$3\kappa^2 = \alpha_1{}^2(1+e_{26}{}^2/c_{66}K_2), \tag{29}$$

where α_1 is the smallest root (>0) of (27).

In the absence of piezoelectric effect ($e_{26}=0$), $\alpha_1=\pi/2$ and $\kappa^2=\pi^2/12$ as in I and II.

PLATE OF FINITE LENGTH

As observed in II, extensive experimental evidence obtained by Sykes on rotated Y-cut quartz plates shows that, over a considerable range, the frequencies are practically independent of the z-dimension. This suggests taking $\zeta=0$ and ξ, η, and φ independent of z.

[3] A. W. Lawson, Phys. Rev. **62**, 71 (1942). Equation (27), is a special case of Lawson's Eq. (57).

Equations (12) and (13) then become

$$\begin{aligned} M_x &= D_1\xi' + F_1\varphi', \\ M_z &= D_2\xi' + F_2\varphi', \\ Q_x &= D_6(\xi + h\eta') + F_6\varphi, \\ P_x &= F_1\xi' - C_1\varphi', \\ P_y &= F_6(\xi + h\eta') - C_2\varphi, \\ M_{xz} &= Q_z = P_z = 0, \end{aligned} \tag{30}$$

where prime indicates differentiation with respect to x. Inserting (30) into (15) (the equations of motion and the electric field equation) the latter become, after introducing a factor $\exp(i\omega t)$,

$$\begin{aligned} D_1\xi'' - D_6(\xi + h\eta') + F_1\varphi'' - F_6\varphi + \rho\omega^2 R\xi = 0, \\ D_6(\xi' + h\eta'') + F_6\varphi' + \rho\omega^2 h\eta = 0, \\ (R + 4\pi C_1)\varphi'' - 4\pi F_1\xi'' = 0. \end{aligned} \tag{31}$$

From the last of (31),

$$\varphi = 4\pi F_1\xi/(R + 4\pi C_1) + Cx + \varphi_0, \tag{32}$$

where C and φ_0 are constants. C may be taken equal to zero, in the present case. The resulting expression for φ is then introduced into the first two of (31), yielding a pair of equations for ξ and η. Eliminating ξ, the equation governing η is found to be

$$(\partial^2/\partial x^2 + \delta_1^2)(\partial^2/\partial x^2 + \delta_2^2)\eta = 0, \tag{33}$$

where

$$\delta_1^2,\ \delta_2^2 = \frac{S\delta_0^4}{2}\{g^* + 1 \pm [(g^* - 1)^2 + 4g^*\bar{\omega}^{*2}/\omega^2]^{\frac{1}{2}}\},$$

and

$$S = D_1/D_6,$$

$$\delta_0^4 = \rho\omega^2/D_1 = \omega^2/(RS\bar{\omega}^2),$$

$$\bar{\omega}^* = \bar{\omega}(1 + N)^{\frac{1}{2}},$$

$$g^* = R/S(1 + J),$$

$$J = \frac{(e_{11} - e_{12}c_{12}/c_{22})^2}{(c_{11} - c_{12}^2/c_{22})(K_1 + e_{12}^2/c_{22})},$$

$$N = \frac{12e_{26}(e_{11} - e_{12}c_{12}/c_{22})}{\pi^2 c_{66}(K_1 + e_{12}^2/c_{22})}.$$

If the electric field is neglected, $J = N = 0$ and δ_0, δ_1 and δ_2 become identical with the corresponding terms in II.

The solution of (33) may be written as

$$\begin{aligned} \eta &= \eta_1 + \eta_2, \\ \eta_1 &= A_1 \sin\delta_1 x + B_1 \cos\delta_1 x, \\ \eta_2 &= A_2 \sin\delta_2 x + B_2 \cos\delta_2 x. \end{aligned} \tag{34}$$

Then, from (31),

$$(1 + N)\xi = (\sigma_1 - 1)h\eta_1' + (\sigma_2 - 1)h\eta_2' + \frac{F_6\varphi_0\bar{\omega}^{*2}}{D_6(\omega^2 - \bar{\omega}^{*2})}, \tag{35}$$

where

$$\sigma_1,\ \sigma_2 = (\delta_1^{-2},\ \delta_2^{-2})S\delta_0^4. \tag{36}$$

The boundary conditions are $M_x = Q_x = M_{xz} = 0$ on $x = \pm l/2$. These may be applied separately to the even and odd modes. For the former, $B_1 = B_2 = 0$ in (34). Then, inserting (34) and (35) in (30), the boundary conditions become

$$A_1(\sigma_1 - 1)\delta_1^2 \sin(\delta_1 l/2) + A_2(\sigma_2 - 1)\delta_2^2 \sin(\delta_2 l/2) = 0, \tag{37}$$

$$A_1\sigma_1\delta_1 \cos(\delta_1 l/2) + A_2\sigma_2\delta_2 \cos(\delta_2 l/2) = \frac{F_6\varphi_0\omega^2}{D_6 h(\bar{\omega}^{*2} - \omega^2)}.$$

Resonance occurs when the determinant of (37) vanishes, i.e., when

$$\sigma_2\delta_1(\sigma_1 - 1)\tan(\delta_1 l/2) = \sigma_1\delta_2(\sigma_2 - 1)\tan\delta_2 l/2. \tag{38}$$

This frequency equation has the same form as that in II and differs only in the definitions of σ_1, σ_2, δ_1 and δ_2. Proceeding as in II, let

$$\gamma = \delta_1 l/2 \tag{39}$$

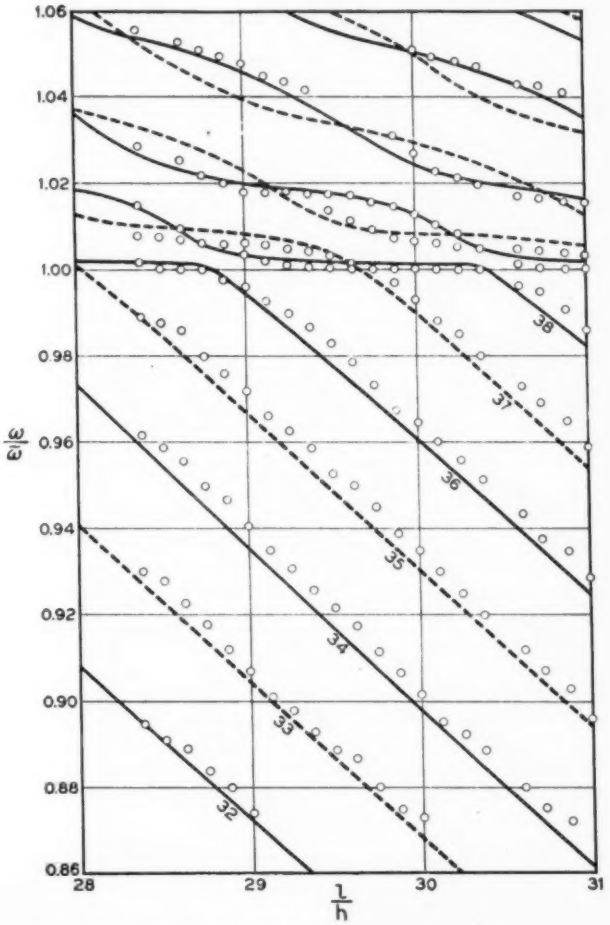


FIG. 2. Comparison of theory with data obtained by Sykes on an AT quartz plate, initially 25×25×0.807 mm, cut down in decrements of 0.1 mm to 25×21.8×0.807 mm.

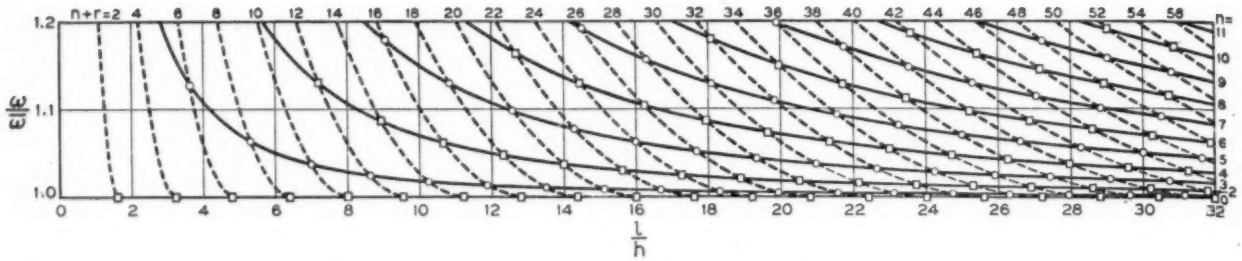


FIG. 3. The circled and squared points are those from the upper half of Fig. 1. The full curves $n=2, 3, 4\cdots$ are those commonly designated as the first, second, third$\cdots$ overtones of the fundamental thickness-shear mode. Only the circled and squared points on these curves correspond to resonant frequencies. The dashed curves, whose equation is (46), are exact at the circled and squared points and, at intermediate points, are approximations to the exact curves in the upper half of Fig. 1.

and

$$a=\delta_2/\delta_1=[(1-b)/(1+b)]^{\frac{1}{2}}, \tag{40}$$

where b is the positive root of

$$b^2=1+4g^*(\bar{\omega}^{*2}/\omega^2-1)(1+g^*)^{-2}. \tag{41}$$

Since δ_1 is always real while δ_2 and a are real if $\omega>\bar{\omega}^*$ and imaginary if $\omega<\bar{\omega}^*$, (38) may be written as

$$\begin{aligned}(a^2-g^*)\tan\gamma&=a(1-g^*a^2)\tan a\gamma, &&\omega>\bar{\omega}^*,\\ (a_1^2+g^*)\tan\gamma&=a_1(1+g^*a_1^2)\tanh a_1\gamma, &&\omega<\bar{\omega}^*,\end{aligned} \tag{42}$$

where $a_1=-ia$. From (40) and (41)

$$\begin{aligned}\omega/\bar{\omega}^*&=[1-a^2(1+g^*)^2/g^*(1+a^2)^2]^{-\frac{1}{2}}, &&\omega>\bar{\omega}^*,\\ \omega/\bar{\omega}^*&=[1+a_1^2(1+g^*)^2/g^*(1-a_1^2)^2]^{-\frac{1}{2}}, &&\omega<\bar{\omega}^*,\end{aligned} \tag{43}$$

and, from (39),

$$\begin{aligned}l/h&=\gamma(\bar{\omega}^*/\omega)[(1+a^2)/3(1+g^*)]^{\frac{1}{2}}, &&\omega>\bar{\omega}^*,\\ l/h&=\gamma(\bar{\omega}^*/\omega)[(1-a_1^2)/3(1+g^*)]^{\frac{1}{2}}, &&\omega<\bar{\omega}^*.\end{aligned} \tag{44}$$

Equations (42), (43), and (44) are identical in form with the correspondingly numbered equations in II and the computation of data for constructing curves of frequency *versus* plate dimensions, for a given crystal and orientation of cut, is carried out in the same manner as before. The only difference is that

$$g=\pi^2c_{66}/12(c_{11}-c_{12}^2/c_{22})$$

of II is replaced here by

$$g^*=g/(1+J),$$

and

$$\bar{\omega}^2=\pi^2c_{66}/\rho h^2$$

of II is replaced here by

$$\bar{\omega}^{*2}=4c_{66}\alpha_1^2(1+e_{26}^2/c_{66}K_2)(1+N)/\rho h^2.$$

For the AT cut of quartz, for example, we find

$$\begin{aligned}g/g^*&=1.000558,\\ \bar{\omega}/\bar{\omega}^*&=1.000552.\end{aligned} \tag{45}$$

It is apparent that it is unnecessary to recompute the curves of II for the new values of g and $\bar{\omega}$. Instead, the computations for the old values were extended to include the odd modes and the results are shown in Figs. 1 and 2 along with additional data by Sykes in the latter figure.

The algebraic formulas for the coordinates of the circled and squared points of Fig. 1 are the same as those given in Eqs. (50) and (51) of II except for the replacement of g and $\bar{\omega}$ by g^* and $\bar{\omega}^*$. Similarly the equations of curves passing through points $n=$constant, $r=$constant and $n+r=$constant are unchanged in form, but the new values for g, $\bar{\omega}$, k, and κ should be used. Figure 4 and Eqs. (52) and (53) of II illustrate the families of curves passing through $n=$constant, and $r=$constant. Figure 3 of the present paper illustrates the curves passing through $n=$constant and $n+r=m=$constant. The equation of the latter is

$$\frac{(1-g)^2}{g}\frac{\omega^2}{\bar{\omega}^2}=\frac{m^2kh^2}{l^2}-2 -\left[\left(\frac{m^2kh^2}{l^2}-2\right)^2-\left(\frac{m^2kh^2}{l^2}\frac{(1-g)}{(1+g)}\right)^2\right]^{\frac{1}{2}}. \tag{46}$$

An interesting feature of (46) is that it gives the frequency spectrum not only exactly at the circled and squared points, but also approximately at intermediate points. This may be seen by comparing the dashed curves in Fig. 3 with the curves in Fig. 1 for $\omega/\bar{\omega}>1$.

The Thermal Dependence of the Elastic Moduli of Polycrystalline Graphite*

F. E. Faris, L. Green, Jr.,† and C. A. Smith
Atomic Energy Research Department, North American Aviation, Inc., Downey, California
(Received September 14, 1951)

Young's modulus and the shear modulus of grade SA-25 molded graphite and grade AUF extruded graphite were measured as functions of temperature in the range from 25°C to 2000°C. In the case of the extruded material, variations of the two moduli with position and direction relative to the axis of extrusion were also investigated. The moduli were found to increase with temperature, a behavior in contrast with that exhibited by most materials. Maxima in the low frequency shear modulus values and increases in the internal friction of both graphites observed in the range from 1550–1750°C may possibly be interpreted in terms of grain boundary relaxation. The bulk modulus and Poisson's ratio at room temperature were calculated from the data for the nominally isotropic SA-25 graphite to be 6×10^{10} dynes/cm² and 0.27, respectively.

I. INTRODUCTION

In contrast with the behavior exhibited by most materials, the short-time tensile strength of various types of commercial graphite has been found to reach a maximum at a temperature of about 2500°C.[1] Because of the unusual temperature dependence of the tensile strength, experiments designed to study the elastic properties over a wide temperature range were undertaken. Some preliminary measurements of Young's modulus already reported[1] showed an increase in the modulus with temperature.

Graphite grades AUF and SA-25‡ were chosen for the investigation as being good examples of oriented and non-oriented graphite, respectively. AUF is an extruded graphite with an orientation of about 60 percent, and SA-25 is a molded material with an orientation of less than one percent and a smaller grain size. As used here the term orientation refers to the relative number of graphite crystals having their basal planes aligned in a certain preferred direction. In extruded graphite the basal planes tend to be aligned in the direction of extrusion.

II. APPARATUS AND PROCEDURE

A. Young's Modulus

Young's modulus was determined as a function of temperature by measuring the frequency of vibration of specimens fabricated as cantilever beams. The beams were 3.25 inches long, 0.25 inch wide, and about 0.050 inch thick and were cut from the extruded graphite rod with their length axes both parallel and perpendicular to the direction of extrusion as shown in Fig. 1 and described in Table I. The specimens were heated in a vacuum by radiation from a surrounding cylinder carrying a direct current as is shown in Fig. 2. The technique of the measurement has been described in detail in reference 1. The room temperature values for Young's modulus were determined by measuring the static deflection of the cantilever specimens under a known load.

B. Shear Modulus

The variation of the shear modulus with temperature was determined from measurements of the frequency of vibration of a torsion pendulum incorporating the graphite sample under study as its suspension. An optical lever utilizing a mirror at the bottom of the pendulum amplified the vibration enough to permit measurement of its logarithmic decrement. The furnace used to heat the samples for the torsion measurements is diagrammed in Fig. 3. The suspensions, one of which is shown in Fig. 4, were machined to diameters of $\frac{1}{32}$ inch and $\frac{3}{16}$ inch for the small and the large sections, respectively, and soldered to a copper bridge at the top of the furnace. Some samples of AUF were cut with their axes in the extrusion direction and some with their axes in a radial direction so that the effect of orientation could be determined.

It was not feasible to obtain absolute values of the shear modulus at room temperature by the method just

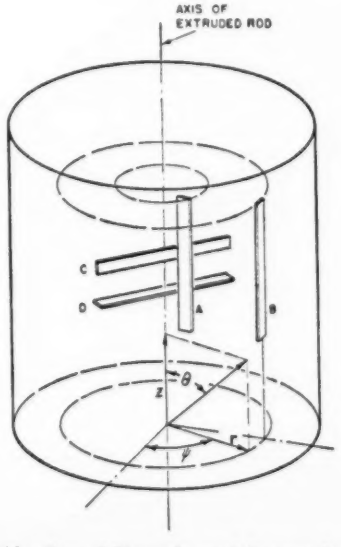


FIG. 1. Identification of Young's modulus specimen orientations.

* The work reported here was done under Contract AT-11-1-GEN-8 with the AEC.
† Present address: Aerojet Engineering Corporation, Azusa, California.
[1] Malmstrom, Keen, and Green, J. Appl. Phys. **22**, 593 (1951).
‡ AUF and SA-25 are types of commercial graphite manufactured by the National Carbon Company.

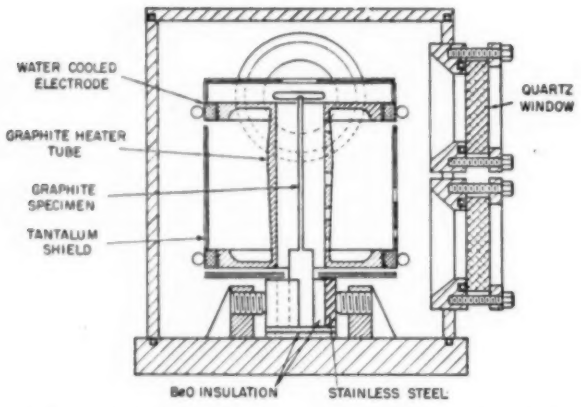


FIG. 2. Cross section of furnace with cantilever beam installed for measurements of Young's modulus.

described because of the small diameter of the specimens. Accordingly, data for the absolute values of the shear modulus at room temperature were obtained by observing the angle of twist induced in $\frac{1}{8}$-inch diameter rods by a known torque.

III. EXPERIMENTAL RESULTS

A. Young's Modulus

Curves showing the temperature dependence of Young's modulus for all samples of AUF are given in Fig. 5 and for all samples of SA-25 in Fig. 6. The temperature variation along the beams amounted to as much as 10 percent in some cases. The temperature values for the curves are averages obtained as described in the appendix. It is estimated that temperatures could be determined with the pyrometer to ± 5°C. Within the limit of error it was found that the results were

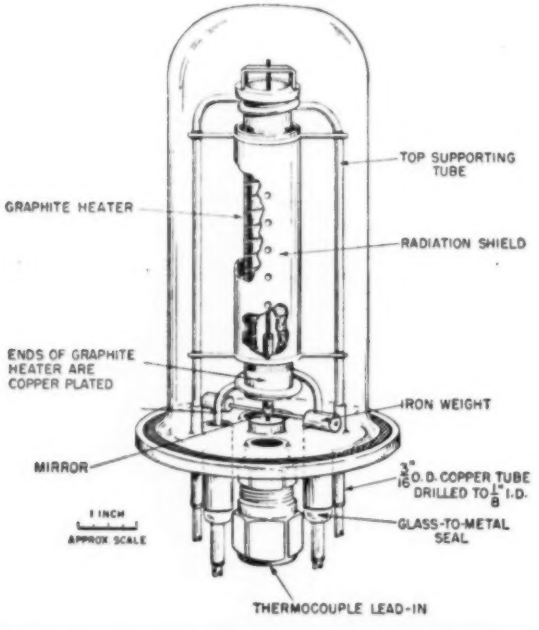


FIG. 3. Furnace used in shear modulus determinations.

independent of the amplitude of vibration. It is to be noted that there is considerable variation from one sample to another. This variation, which is inherent in graphite experiments, arises from the lack of homogeneity and isotropy of the material. A typical density profile across an extruded rod of AUF graphite is presented in Fig. 7.

In order to determine how well the results could be reproduced from run to run, one sample was measured twice. It was found that within the experimental error there was no difference between the two runs.

B. The Shear Modulus

Data for $(\omega_T/\omega_0)^2$, in which ω_T is the circular frequency at a temperature T and ω_0 is the circular frequency at room temperature, are plotted as a function of the maximum temperature along the sample in Figs. 8, 9, and 10 for AUF graphite cut parallel to the direction of extrusion, AUF cut along a radial direction, and SA-25, respectively. The standard deviation in the values of $(\omega_T/\omega_0)^2$ is in most caes about one percent based on four determinations at each temperature. The deviation is somewhat greater in the values for parallel-cut AUF near 2000°C. The data for parallel-cut AUF

TABLE I. Identification of specimen orientations (see Fig. 1).

Group designation	Location relative to axis of rod	Orientation of specimen length axis relative to extrusion direction	Average density (g/cm³)	Plane of vibration
A	Within 1-in. radius	Parallel	1.545	Not controlled
B	On 5-in. radius	Parallel	1.610	ψ=constant
C	Within 2-in. radius	Perpendicular	...	z=constant
D	Within 2-in. radius	Perpendicular	...	ψ=constant

were the first taken, and at the higher temperatures only one or two frequency measurements were made at each temperature.

Although the top support was water cooled, its underside received a large amount of radiant heat when the furnace was at high temperatures. It was therefore decided to conduct an experiment to make sure that changes in the solder joint were not affecting the results. The temperature of the top support was measured with a thermocouple when the furnace was at its maximum temperature. The frequency of a sample was then measured, first at room temperature in the usual way, and again at room temperature but with the support heated to the same temperature as that observed when the furnace was on. No difference in frequency was found, thus showing that no error was being introduced by the joint.

In order to find out how well the results could be reproduced, one sample each of SA-25 and parallel-cut AUF were measured twice. The sample of AUF was actually installed twice; whereas measurements on the SA-25 sample were made with the same installation, first with increasing temperature values and then with de-

creasing. As may be seen from the points in Figs. 8 and 10, the reproducibility was satisfactory. It is to be noted that for one measurement of the AUF sample (sample 2 in Fig. 8) the room-temperature frequency was 1.28 cycles/sec, and for the other measurement of the same sample it was 1.74 cycles/sec. It thus appears that a small change in the room-temperature frequency has no significant effect upon the variation with temperature.

The torsion bar and weights normally used in the experiment had a mass of 17.8 grams, which resulted in a tensile stress of about 35×10^5 dynes/cm² in the sample. It was decided to measure one sample with smaller weights on the bar to determine the effect of changing the stress. Accordingly, the weights were

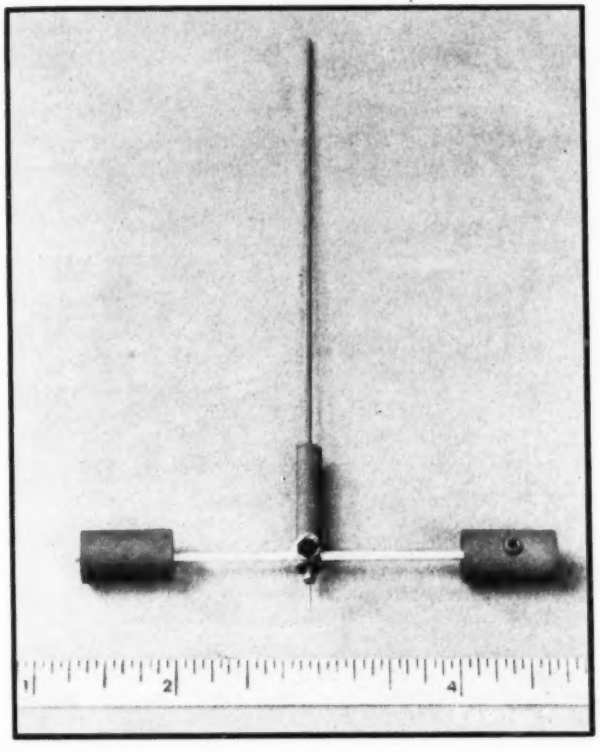


FIG. 4. Graphite suspension and torsion pendulum attachments for measurements of the relative values of the modulus of rigidity.

changed so that the mass supported by the suspension was 3.5 grams, corresponding to a stress of about 6.9×10^5 dynes/cm². It was found that the points for the smaller stress lay about one percent above those for the greater; however, all points depended on the room-temperature frequency, and this frequency might well have been in error by one percent. It is thus considered that for tensile stress levels below 35×10^5 dynes/cm² the effect of changing the tensile stress upon the variation of the modulus with temperature is quite small.

The maximum shear strain during the observations was of the order of 10^{-4}, which corresponds to a maximum shear stress in the outer fibers of about 30×10^5 dynes/cm². Measurements of frequency starting with

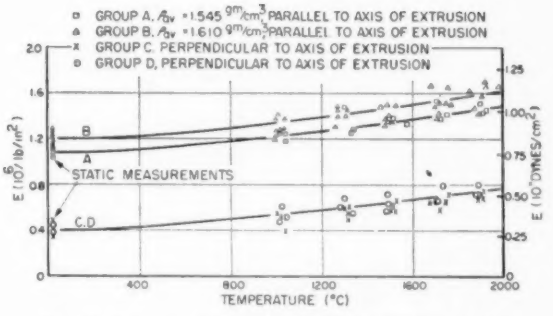

GROUP A. ρ_{av} = 1.545 gm/cm³ PARALLEL TO AXIS OF EXTRUSION
GROUP B. ρ_{av} = 1.610 gm/cm³ PARALLEL TO AXIS OF EXTRUSION
GROUP C. PERPENDICULAR TO AXIS OF EXTRUSION
GROUP D. PERPENDICULAR TO AXIS OF EXTRUSION
STATIC MEASUREMENTS
E (10⁶ lb/in²)
E (10¹⁰ DYNES/cm²)
TEMPERATURE (°C)

FIG. 5. Temperature variation of Young's modulus for AUF graphite.

amplitudes corresponding to stresses less than half of this value gave the same value for the frequency to within the experimental error; thus one can state that within the accuracy of measurement the modulus values were independent of amplitude.

It was found that the internal friction was not amplitude-independent, but compensation for this effect insofar as temperature dependence is concerned was provided to some extent by always measuring the damping for the same amplitude range. It was also found that measurements of the internal friction could not be reproduced well, and the difficulties were traced to friction in the bearing which constrained the motion to rotation. There were certain general features of the temperature dependence which were consistent, however, and for this reason the curve obtained for a sample of SA-25 graphite is shown in Fig. 10. The quantity chosen as a measure of the internal friction is δ/π, where δ is the logarithmic decrement.[2] Although the absolute values in Fig. 10 are undoubtedly high, the curve does illustrate very well the trend observed for all samples.

Because of the furnace design there was a more severe temperature gradient along the samples than in the case of the measurements of Young's modulus. Typical temperature distributions are shown in Fig. 11. From the known temperature distributions the ratio (G_T/G_0),

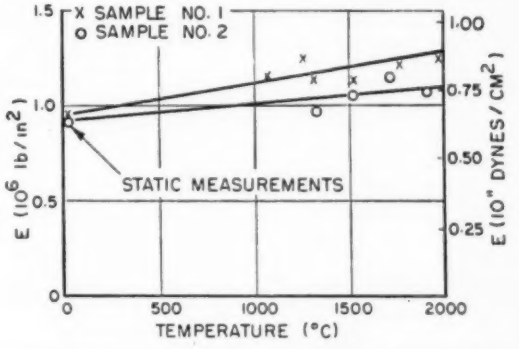


FIG. 6. Temperature variation of Young's modulus for SA-25 graphite.

[2] C. Zener, *Elasticity and Anelasticity of Metals* (The University of Chicago Press, Chicago, Illinois, 1948), p. 61.

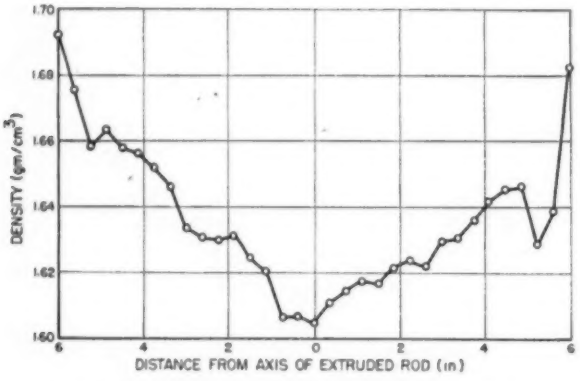


FIG. 7. Typical density profile of extruded AUF rod.

where G_T is the modulus at a temperature T and G_0 is the value at room temperature, was obtained as a function of temperature by the method outlined in the appendix. The results are shown in Fig. 12. The curve shown for each type of graphite represents an average of the results for all samples of a given kind. The temperature dependence for AUF cut parallel to the extrusion direction and for AUF cut in a radial direction is nearly the same, and hence only one curve has been drawn to cover both cases. On the basis of the experimental error and the approximations made in correcting for the temperature distribution, it is estimated that the (G_T/G_0) values may have an error of 5 percent associated with them.

The estimate of 5 percent is of course for the average value of the samples measured. It is to be noted that different specimens of the same type of graphite varied appreciably from one another; hence measurements on other samples might well have led to somewhat different average values. In particular it is quite likely that the percentage increase with temperature for the samples cut in a radial direction should be somewhat greater than for the samples cut parallel to the direction of extrusion.

The absolute values for the modulus of rigidity of the $\frac{1}{8}$-inch diameter rods and for Young's modulus of the same rods at room temperature under static conditions

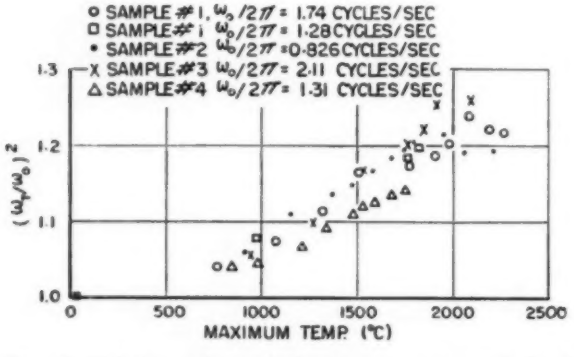

o SAMPLE #1, ω₀/2π = 1.74 CYCLES/SEC
□ SAMPLE #1 ω₀/2π = 1.28 CYCLES/SEC
• SAMPLE #2 ω₀/2π = 0.826 CYCLES/SEC
X SAMPLE #3 ω₀/2π = 2.11 CYCLES/SEC
△ SAMPLE #4 ω₀/2π = 1.31 CYCLES/SEC
(ωT/ω0)²
1.3
1.2
1.1
1.0
0 500 1000 1500 2000 2500
MAXIMUM TEMP. (°C)

FIG. 8. Relative values of the frequency squared (torsion samples) as a function of maximum temperature along the rod for AUF graphite parallel to the direction of extrusion.

are shown together with their standard deviations in Table II. There should be very little difference between the room-temperature values measured statically and those measured dynamically at the frequencies chosen for the experiment.

IV. DISCUSSION OF THE RESULTS

Since SA-25 graphite is nearly isotropic in a macroscopic sense, a knowledge of the shear modulus and Young's modulus may be used to calculate values for Poisson's ratio and the bulk modulus. The necessary relationships are given in a number of places, for example in reference 3. Values for Poisson's ratio μ and the bulk modulus k at room temperature are given in Table II. A small change in Young's modulus E or the shear modulus G results in a large change in k and μ, and as a result the values for the two samples differ rather widely.

The increase in the moduli as the temperature is increased above 25°C is in contrast with the behavior of most materials. In so far as plausible mechanisms to

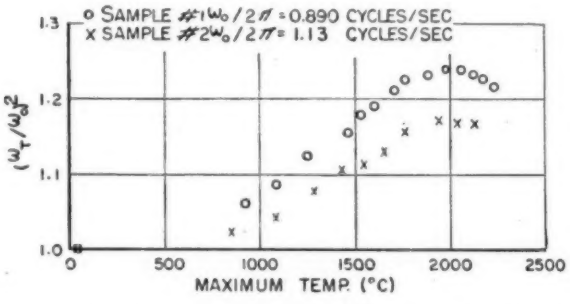


FIG. 9. Relative values of the frequency squared (torsion samples) as a function of maximum temperature along the rod for AUF graphite cut in a radial direction.

account for the increase are concerned, only negative statements can be made at present. On the basis of measurements repeated on the same samples for both moduli and measurements made on one torsion sample over a period of 15 minutes while the sample was maintained at a temperature near 2000°C, one can state that the effect is reversible and time-independent. It would thus appear that mechanisms such as the diffusion of carbon atoms between grains or the reorientation of grains are ruled out. If one attempts to explain the increase on the basis of an increase in the binding forces within the crystals, the large percentage increase in Young's modulus for AUF samples cut in a radial direction as compared to that for AUF samples cut in a longitudinal direction would seem to require a large increase in the binding between planes. Graphite crystals have a large positive temperature coefficient of expansion perpendicular to the basal plane,[4] and the

[3] I. S. Sokolnikoff, *Mathematical Theory of Elasticity* (McGraw-Hill Book Company, Inc., New York, 1946), pp. 68 and 70.

[4] J. B. Nelson and D. P. Riley, Proc. Phys. Soc. (London) **57**, 477 (1945).

distance between planes is so large that it is difficult to conceive of a change in the nature of the binding force between planes. These facts appear to rule out any increase in the binding between planes with an increase in temperature. It would, of course, be interesting to have data on the temperature dependence of the elastic coefficients of a graphite crystal in order to check this surmise.

A possible explanation of the decline which was observed in the shear modulus after sufficiently high temperatures were reached may lie in the mechanism of grain boundary relaxation. Because of the porous nature of artificial graphite there is some doubt as to whether one can draw a parallel between grain boundary relaxation in metals and in graphite. It is, however, interesting to observe that if grain boundary relaxation is accepted as the mechanism responsible for the decline in the modulus at high temperature, a certain analogy does exist between the behavior of metals and that of graphite. Ke[5] has shown that in the case of a polycrystalline metal the internal friction and the relaxation

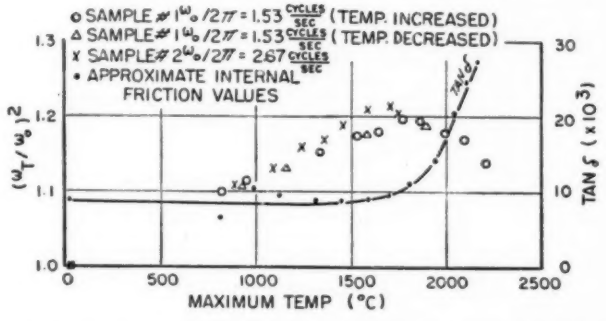

o SAMPLE # $\omega_0/2\pi$ = 1.53 CYCLES/SEC (TEMP. INCREASED)
Δ SAMPLE # $\omega_0/2\pi$ = 1.53 CYCLES/SEC (TEMP. DECREASED)
x SAMPLE # $2\omega_0/2\pi$ = 2.67 CYCLES/SEC
• APPROXIMATE INTERNAL FRICTION VALUES
$(\omega_T/\omega_0)^2$
TAN δ (x10³)
MAXIMUM TEMP (°C)

FIG. 10. Relative values of the frequency squared (torsion samples) as a function of maximum temperature along the rod for SA-25 graphite.

in the modulus associated with grain boundary movement may be expressed as a function of the parameter α with

$$\alpha = g\nu e^{H/RT},$$

where g is the grain size, ν is the frequency, H is the heat of activation, R is the gas constant, and T is the absolute temperature. From the information available[6,7] it is estimated that the grain size of SA-25 is about one-tenth that of AUF. A smaller grain size is consistent with the fact that the modulus of SA-25 relaxes at a lower temperature than that of AUF. If it is assumed that the decrease in the modulus is the result of grain-boundary relaxation and that the process may be described by the parameter α, an order-of-magnitude value for the activation energy may be calculated. The activation energy is found to be roughly 6×10^4

[5] T. Ke, Phys. Rev. **72**, 41 (1947).
[6] H. L. Larson, National Carbon Company, Inc., Cleveland, Ohio, letter of September 15, 1947, to H. Pearlman.
[7] H. G. MacPherson, National Carbon Company, Inc., Cleveland, Ohio, private communication of October, 1948, to H. Pearlman.

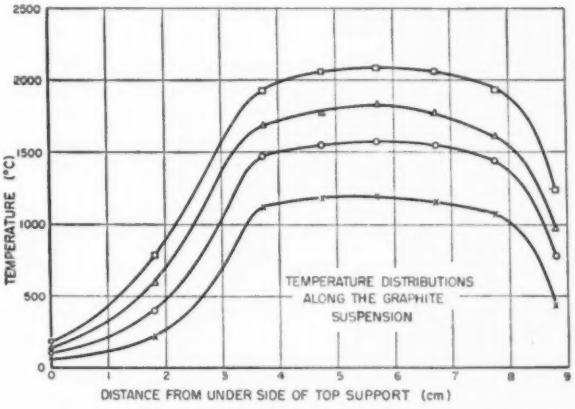


FIG. 11. Some typical temperature distributions for the torsion samples.

cal/mole. Using this value for the activation energy one finds that for frequencies of about 70 cycles/sec the relaxation in the modulus would occur above 2000°C for both AUF graphite and SA-25 graphite. Because of the frequencies used, maxima in the curves for Young's modulus should therefore not be expected. It should be noted that the small changes in frequency from sample to sample in the measurements of the modulus of rigidity would shift the relaxation portion of the curves very little.

From Fig. 5 the effect of density upon Young's modulus may be observed. Although there is considerable scatter in the results, it is apparent that the modulus tends to increase as the density increases. Since a larger density indicates less porosity, the trend is in the expected direction.

Also in Fig. 5 it is to be observed that there is little if any differences in Young's modulus for radial samples which vibrated in a plane perpendicular to the extrusion axis as compared to Young's modulus for radial samples which vibrated in a plane defined by the axis and a diameter. This result indicates that the graphite crystals tend to be aligned with their basal planes perpendicular to the radial direction. This type of alignment is to be expected as a result of the forces present during the extrusion process used to shape AUF graphite.

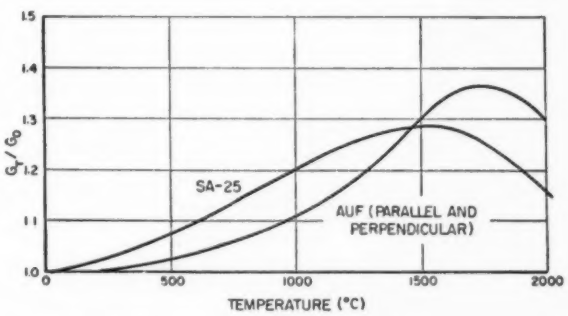


FIG. 12. Ratio of the modulus of rigidity at a temperature T to its value at room temperature as a function of temperature.

TABLE II. Room temperature values of the modulus of rigidity G, Young's modulus E, the bulk modulus k, and Poisson's ratio μ for the torsion specimens. The limits quoted are the standard deviations for individual samples.

Material	Sample	$G\times10^{-10}$ (dynes/cm²)	$E\times10^{-10}$ (dynes/cm²)	$k\times10^{-10}$ (dynes/cm²)	μ
AUF Parallel	1	3.05±0.13	8.56±0.30	...	...
	2	3.10±0.13	9.53±0.30	...	...
AUF[a] Radial	1	2.43±0.10	4.31±0.17	...	...
	2	2.40±0.10	4.58±0.19	...	...
SA-25	1	3.32±0.14	8.08±0.30	4.7±1.1	0.21±0.07
	2	3.21±0.14	8.56±0.30	8.5±4.1	0.33±0.08

[a] Because of the relatively long length of the specimens (approximately 8½ in.) the radial samples were not uniform in density or orientation.

V. CONCLUSIONS

The experiment has produced average values for the modulus of rigidity and Young's modulus of SA-25 graphite and AUF graphite cut parallel to the extrusion direction and in a radial direction in the temperature range from 25°C to 2000°C. The increase which occurs is time-independent and reversible. The data show that Young's modulus increases with density and that Young's modulus for AUF samples cut in a radial direction is independent of the plane of vibration.

It has been shown that certain mechanisms which might give rise to an increase in the moduli, namely, diffusion of carbon atoms between grains, reorientation of grains, and an increase in binding within the crystals are apparently eliminated by the experimental results. It has been suggested that the decrease in the modulus of rigidity which was observed at relatively high temperatures may be the result of grain boundary relaxation.

Average values for Poisson's ratio and the bulk modulus for SA-25 graphite at room temperature have been calculated to be 0.27 and 6×10^{10} dynes/cm², respectively.

ACKNOWLEDGMENTS

The authors wish to thank Mr. Carl Malmstrom and Dr. Sidney Siegel for their helpful suggestions concerning the paper. It is also a pleasure to acknowledge the help of Mr. Arthur Miller, Mr. Harry Strahl, and Mr. James McClelland in obtaining some of the data, and the painstaking machine work of Mr. Sidney Becker and Mr. Patrick Nagle, who made the graphite samples.

APPENDIX

Corrections for the temperature distribution along the beams used for the measurement of Young's modulus are made by determining an average temperature for each modulus value, the average being calculated by using the nominal strain in the outer fibers of the specimen as a weight function. This method of correction is justified because of the relatively small temperature gradient along the sample, the maximum observed temperature variation being about ten percent, and because of the mild functional dependence of the modulus upon the temperature.

The relatively large temperature variation which existed in the case of the torsion samples necessitates a more detailed correction than in the case of the beams.

In making the correction for the torsion samples, we must first determine what quantity the square of the frequency represents when the modulus is a function of the distance along the suspension. Let ω be the circular frequency so that

$$\omega^2=L'/I, \tag{1}$$

where L' is the torque per unit angle turned through by the torsion rod, and I is the moment of inertia of the torsion rod and weights. Now consider the situation for a certain angle of twist θ_0 at the end of the suspension. If $G(z)$ and $\theta(z)$ are the shear modulus and the angle of twist at a distance z from the support, we have

$$S=G(z)\rho(d\theta/dz), \tag{2}$$

where S is the stress at the circumference and ρ is the radius of the suspension. Since the internal moment must balance the applied torque, we also have the relationship

$$S=L\rho/J, \tag{3}$$

where L is the torque and J is the polar moment of inertia. From Eq. (2) we obtain, since S is constant for a given L,

$$S/\rho\int_0^l dz/G(z)=\theta_0, \tag{4}$$

in which l is the length of the suspension. Eliminating S between (3) and (4) gives

$$\frac{L}{\theta_0}=L'=\frac{J}{\int_0^l[1/G(z)]dz}. \tag{5}$$

Substituting in (1) results in the relationship

$$\omega^2=\frac{J}{Il\langle1/G(z)\rangle_{Av}}, \tag{6}$$

where $\langle1/G(z)\rangle_{Av}$ represents the average value of $1/G(z)$ defined in the usual way by

$$\langle1/G(z)\rangle_{Av}=1/l\int_0^l dz/G(z).$$

It is thus apparent from (6) that an experimental determination of ω^2 for a suspension whose modulus of rigidity is not constant is inversely proportional to the average value of the reciprocal of G.

In order to apply the information just obtained to the correction of the data we proceed as follows. If ω_0 is the circular frequency, and G_0 is the value of the modulus, both taken at room temperature, we have

$$\langle1/G(z)\rangle_{Av}=\omega_0^2/G_0\omega^2, \tag{7}$$

in which the average value is for the temperature distribution present when ω was observed. Implicit in expression (7) is the assumption that J/l is a constant regardless of temperature. Actually the ratio varies because of thermal expansion so that (7) implies a redefinition of the modulus G at some temperature T in terms of the dimensions at room temperature rather than at the temperature T. The matter is of little practical importance here because measurements of the increase in length for the conditions of the experiment showed that the maximum change in J/l was of the order of one percent. Now let G_0/G_T in which G_0 and G_T are the values of the modulus at room temperature and at a temperature T, respectively, be represented by a polynomial in powers of $(T-T_0)$, adequate to represent the data over a restricted interval. We then have

$$G_0/G_T=1+A_1(T-T_0)+A_2(T-T_0)^2 +A_3(T-T_0)^3+\cdots, \quad (8)$$

with the coefficients A_n to be determined from the experimental data. The determination is made by solving a set of simultaneous equations involving the A_n. A particular equation of the set is obtained by substituting an experimentally observed value for $(\omega_0/\omega)^2$ in Eq. (7) and by substituting for $\langle G_0/G(z)\rangle_{Av}$ the reciprocal of the rod length multiplied by the integral of Eq. (8) over the length of the rod. The integration is performed by using the temperature distribution which was observed for the value of $(\omega_0/\omega)^2$ selected. In actual practice the corrected curve was obtained by determining three coefficients which matched the integrated experimental results for three points at 1500°C and above and by determining three coefficients which matched the results for three points at 1500°C and below. The corrected curve for the high region and that for the low region were observed to agree well in an intermediate range, and the final corrected curve was obtained by a smooth joining of the low temperature and the high temperature curves.

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 1 JANUARY, 1952

A Nonsteady Heat Diffusion Problem with Spherical Symmetry*

M. S. Plesset and S. A. Zwick
California Institute of Technology, Pasadena, California
(Received August 30, 1951)

A solution in successive approximations is presented for the heat diffusion across a spherical boundary with radial motion. The approximation procedure converges rapidly provided the temperature variations are appreciable only in a thin layer adjacent to the spherical boundary. An explicit solution for the temperature field is given in the zero order when the temperature at infinity and the temperature gradient at the spherical boundary are specified. The first-order correction for the temperature field may also be found. It may be noted that the requirements for rapid convergence of the approximate solution are satisfied for the particular problem of the growth or collapse of a spherical vapor bubble in a liquid when the translational motion of the bubble is neglected.

I. INTRODUCTION

A PROBLEM of nonsteady heat diffusion is encountered if one considers the dynamics of a vapor bubble in a liquid. As the size of the bubble changes, heat flows across the moving liquid-vapor interface. The liquid is assumed here to be nonviscous and incompressible, and the thermal conductivity k, density ρ, and specific heat c of the liquid are assumed to have insignificant variation with temperature. The temperature T in the liquid then satisfies the equation

$$\Delta T=\frac{1}{D}\frac{dT}{dt}-\frac{1}{k}\dot{\eta}. \quad (1)$$

Here $D=k/\rho c$ is the thermal diffusivity of the liquid, $\eta=\eta(t)$ is the heat source per unit volume in the liquid which is taken to be a function of time only, and dT/dt denotes the particle derivative so that

$$dT/dt=\partial T/\partial t+\mathbf{v}\cdot\nabla T,$$

where $\mathbf{v}$ is the liquid velocity which in general varies with position and time.

II. FORMULATION OF THE PROBLEM

It is advantageous to transform Eq. (1) from Eulerian to Lagrangian coordinates. It will be assumed that the motion possesses spherical symmetry; i.e., the vapor bubble is spherical, and its radial motion is sufficiently rapid that any translational motion may be neglected. The Eulerian coordinates will be chosen as (r, t) with origin $r=0$ at the center of the bubble. If $R(t)$ is the bubble radius at time t, Lagrange coordinates may be defined by

$$h=(1/3)[r^3-R^3(t)], \quad (2a)$$

$$l=t, \quad (2b)$$

* This study was supported by the ONR.

and Eq. (1) can now be written as

$$\frac{\partial}{\partial h}\left(r^4\frac{\partial T}{\partial h}\right)=\frac{1}{D}\frac{\partial T}{\partial t}-\frac{1}{k}\dot{\eta}. \tag{3}$$

If at $t=0$, the temperature at $r=\infty$ is T_0, then at a later time t the temperature at infinity is $T_\infty=T_0+(D/k)\eta(t)$, where $\eta(0)=0$. It is convenient to introduce a function $U(h, t)$ such that

$$\partial U/\partial h=T-T_0, \tag{4}$$

and Eq. (3) may then be written

$$\frac{\partial}{\partial h}\left\{r^4\frac{\partial^2 U}{\partial h^2}-\frac{1}{D}\frac{\partial U}{\partial t}\right\}=-\frac{1}{k}\dot{\eta},$$

so that by integration

$$r^4\frac{\partial^2 U}{\partial h^2}-\frac{1}{D}\frac{\partial U}{\partial t}=-\frac{h}{k}\dot{\eta}+\psi(t), \tag{5}$$

where $\psi(t)$ is an arbitrary function of time. From Eq. (4),

$$U=\int_0^h (T-T_0)dh+\chi(t),$$

and the function $\chi(t)$ may be chosen so that $\psi(t)=0$, and also so that

$$U(h, 0)=0, \tag{6}$$

if it is assumed that the temperature at $t=0$ is T_0 everywhere. Equation (5) now becomes

$$r^4\frac{\partial^2 U}{\partial h^2}-\frac{1}{D}\frac{\partial U}{\partial t}=-\frac{h}{k}\dot{\eta}. \tag{7}$$

In addition to Eq. (6), one has the boundary condition at infinity

$$(\partial U/\partial h)_{h=\infty}=(D/k)\eta(t). \tag{8}$$

The further boundary condition required to fix the temperature field may be given by the temperature gradient at the bubble wall, $r=R(t)$, as a function of time. One then has

$$(\partial^2 U/\partial h^2)_{h=0}=(1/R^2)(\partial T/\partial r)_{r=R(t)}=F(t). \tag{9}$$

The diffusion problem thus defined can be solved by successive approximations, if the assumption is made that U and T vary appreciably only in the region

$$h\ll R^3(t). \tag{10}$$

Physically, for the vapor bubble problem, this assumption is made plausible by the fact that not only is the heat capacity much greater in the liquid state than in the vapor state but the thermal diffusivity is also much smaller.

III. ZERO-ORDER SOLUTION

It is convenient to introduce in Eq. (7) a new time variable τ defined by

$$\tau=\int_0^t R^4(t)dt.$$

Equation (7) now becomes, to the first order in h/R^3,

$$\frac{\partial^2 U}{\partial h^2}-\frac{1}{D}\frac{\partial U}{\partial \tau}+\frac{h}{k}\frac{d\eta}{d\tau}=-\frac{4h}{R^3}\frac{\partial^2 U}{\partial h^2}, \tag{11}$$

since

$$r^4\approx R^4(1+4h/R^3)$$

to this order. By the usual procedure of successive approximations, one sets

$$U=U^0+U'+\cdots, \quad T=T^0+T'+\cdots,$$

where the superscript denotes the order of the approximation in powers of the perturbation parameter h/R^3. The zero-order approximation, U^0, is thus determined by

$$\frac{\partial^2 U^0}{\partial h^2}-\frac{1}{D}\frac{\partial U^0}{\partial \tau}+\frac{h}{k}\frac{d\eta}{d\tau}=0, \tag{12}$$

with the boundary conditions (6), (8), and (9).

The appropriate solution of Eq. (12) is readily found by taking the Laplace transform on the variable τ. If

$$u(h, s)=\mathfrak{L}\{U^0(h, \tau)\}\equiv\int_0^\infty e^{-s\tau}U^0(h, \tau)d\tau,$$

$$\sigma(s)=\mathfrak{L}\{\eta(\tau)\},$$

$$f(s)=\mathfrak{L}\{F(\tau)\},$$

then

$$\frac{d^2u}{dh^2}-\frac{s}{D}u=-\frac{hs}{k}\sigma(s), \tag{13}$$

where use has been made of Eq. (6) and the specification that $\eta(0)=0$. One has further, from Eqs. (8) and (9),

$$(du/dh)_{h=\infty}=(D/k)\sigma(s), \tag{14}$$

$$(d^2u/dh^2)_{h=0}=f(s). \tag{15}$$

The required solution of Eq. (13) is

$$u=(D/s)f(s)\exp\{-h(s/D)^{\frac{1}{2}}\}+h(D/k)\sigma(s),$$

so that

$$\frac{du}{dh}=-(D/s)^{\frac{1}{2}}f(s)\exp\{-h(s/D)^{\frac{1}{2}}\}+(D/k)\sigma(s). \tag{16}$$

From Eq. (16), one finds

$$\mathfrak{L}^{-1}\{du/dh\}=\partial U^0/\partial h=T^0(h,\tau)-T_0$$

$$=(D/k)\eta(\tau)-(D/\pi)^{\frac{1}{2}}\int_0^\tau \frac{F(\zeta)}{(\tau-\zeta)^{\frac{1}{2}}}\times\exp\left\{-\frac{h^2}{4D(\tau-\zeta)}\right\}d\zeta. \quad (17)$$

If one sets

$$\theta(h,\tau)=T-T_0-(D/k)\eta(\tau)=T-T_\infty,$$

then

$$\theta^0(h,\tau)=-(D/\pi)^{\frac{1}{2}}\int_0^\tau \frac{F(\zeta)}{(\tau-\zeta)^{\frac{1}{2}}}\times\exp\left\{-\frac{h^2}{4D(\tau-\zeta)}\right\}d\zeta. \quad (17')$$

Thus, the difference between the temperature at the spherical boundary, $T^0(0,\tau)$, and the temperature at infinity, T_∞, is given by

$$\theta^0(0,\tau)=-(D/\pi)^{\frac{1}{2}}\int_0^\tau \frac{F(\zeta)d\zeta}{(\tau-\zeta)^{\frac{1}{2}}}. \quad (18)$$

One also has from Eq. (16) for $h=0$

$$\theta^0(0,\tau)=\mathfrak{L}^{-1}\{(D/s)^{\frac{1}{2}}f(s)\},$$

so that the inverse Laplace transform of Eq. (16) may be written in an alternative form as

$$\theta^0(h,\tau)=[h/(4\pi D)^{\frac{1}{2}}]\int_0^\tau \frac{\theta^0(0,\zeta)}{(\tau-\zeta)^{\frac{3}{2}}}\times\exp\left\{-\frac{h^2}{4D(\tau-\zeta)}\right\}d\zeta. \quad (19)$$

In terms of the original time variable, one has, for example, from Eq. (18)

$$T^0(0,t)-T_0=(D/k)\eta(t)-(D/\pi)^{\frac{1}{2}}\int_0^t \frac{R^2(x)(\partial T/\partial r)_{r=R(x)}}{\left\{\int_x^t R^4(y)dy\right\}^{\frac{1}{2}}}dx. \quad (20)$$

If the variations in $R(t)$ are sufficiently small, Eq. (20) simplifies to

$$T^0(0,t)-T_0=(D/k)\eta(t)-(D/\pi)^{\frac{1}{2}}\int_0^t \frac{(\partial T/\partial r)_{r=R(x)}}{(t-x)^{\frac{1}{2}}}dx, \quad (21)$$

which represents the "plane approximation" obtained if the curvature of the boundary $r=R(t)$ is completely neglected.

It may be noted, when

$$T^0(0,\tau)-T_0-(D/k)\eta(\tau)=\theta^0(0,\tau)$$

is a monotonic function of τ, that one obtains from Eq. (19) the inequality

$$\frac{\theta^0(h,\tau)}{\theta^0(0,\tau)}\leq erfc\{h/(4D\tau)^{\frac{1}{2}}\}.$$

IV. FIRST-ORDER CORRECTION TO THE SOLUTION

If one continues the procedure of successive approximation, the right side of Eq. (11) is now considered as determined from the zero-order solution U^0, and the first-order correction U' is determined by

$$\frac{\partial^2 U'}{\partial h^2}-\frac{1}{D}\frac{\partial U'}{\partial \tau}=-\frac{4h}{R^3}\frac{\partial^2 U^0}{\partial h^2}=G(h,\tau). \quad (22)$$

The boundary conditions for $U'(h,\tau)$ are

$$U'(h,0)=(\partial U'/\partial h)_{h=\infty}=(\partial^2 U'/\partial h^2)_{h=0}=0. \quad (23)$$

From Eq. (4), the first-order correction, T', to the temperature field is given by

$$T'=\partial U'/\partial h.$$

Denoting the Laplace transforms of $U'(h,\tau)$, $G(h,\tau)$ by $v(h,s)$, $g(h,s)$, respectively, one has

$$\frac{d^2 v}{dh^2}-\frac{s}{D}v=g(h,s), \quad (24)$$

with the boundary conditions

$$(dv/dh)_{h=\infty}=(d^2v/dh^2)_{h=0}=0. \quad (25)$$

By Eq. (22), $G(0,\tau)=0$, and the solution to Eqs. (24) and (25) is readily found to be

$$v(h,s)=-\frac{1}{2}\left(\frac{D}{s}\right)^{\frac{1}{2}}\left\{e^{h(s/D)^{\frac{1}{2}}}\int_h^\infty e^{-x(s/D)^{\frac{1}{2}}}g(x,s)dx+e^{-h(s/D)^{\frac{1}{2}}}\int_0^h e^{x(s/D)^{\frac{1}{2}}}g(x,s)dx-e^{-h(s/D)^{\frac{1}{2}}}\int_0^\infty e^{-x(s/D)^{\frac{1}{2}}}g(x,s)dx\right\},$$

so that

$$\frac{dv}{dh}=-\frac{1}{2}\left\{e^{h(s/D)^{\frac{1}{2}}}\int_h^\infty e^{-x(s/D)^{\frac{1}{2}}}g(x,s)dx-e^{-h(s/D)^{\frac{1}{2}}}\int_0^h e^{x(s/D)^{\frac{1}{2}}}g(x,s)dx+e^{-h(s/D)^{\frac{1}{2}}}\int_0^\infty e^{-x(s/D)^{\frac{1}{2}}}g(x,s)dx\right\}. \quad (26)$$

For $h=0$, one has

$$\left(\frac{dv}{dh}\right)_{h=0} = -\int_0^\infty e^{-x(s/D)^{\frac{1}{2}}} g(x,s)dx = \mathfrak{L}\{T'(0,\tau)\}. \quad (27)$$

From Eq. (27), the first-order correction to the temperature at the spherical boundary is

$$T'(0,\tau) = -\int_0^\infty dx \int_0^\tau G(x,\xi) \times \frac{x\exp[-x^2/4D(\tau-\xi)]}{[4\pi D(\tau-\xi)^3]^{\frac{1}{2}}} d\xi. \quad (28)$$

By Eq. (22) and the definition of θ^0

$$G(h,\tau) = -\frac{4h}{R^3(\tau)} \frac{\partial\theta^0(h,\tau)}{\partial h},$$

so that Eq. (28) becomes

$$T'(0,\tau) = \int_0^\infty dx \int_0^\tau \frac{4x^2}{R^3(\xi)} \frac{\partial\theta^0(x,\xi)}{\partial x} \times \frac{\exp[-x^2/4D(\tau-\xi)]}{[4\pi D(\tau-\xi)^3]^{\frac{1}{2}}} d\xi,$$

and, if it is permissible to interchange the order of the integrations, one gets

$$T'(0,\tau) = 2\int_0^\tau \frac{d\xi}{R^3(\xi)[\pi D(\tau-\xi)^3]^{\frac{1}{2}}} \int_0^\infty \frac{\partial\theta^0(x,\xi)}{\partial x} \times \exp[-x^2/4D(\tau-\xi)]x^2dx. \quad (29)$$

Differentiation of Eq. (19) gives

$$\frac{\partial\theta^0(x,\xi)}{\partial x} = \frac{1}{(4\pi D)^{\frac{1}{2}}} \int_0^\xi \frac{\theta^0(0,\zeta)}{(\xi-\zeta)^{\frac{1}{2}}} \frac{\partial}{\partial x} \times \{x\exp[-x^2/4D(\xi-\zeta)]\}d\zeta.$$

Substitution of this relation in Eq. (29) leads to the result

$$T'(0,\tau) = 2\left(\frac{D}{\pi}\right)^{\frac{1}{2}} \int_0^\tau \frac{d\xi}{R^3(\xi)} \times \int_0^\xi \frac{\theta^0(0,\zeta)}{(\tau-\zeta)^{\frac{1}{2}}} \left[1-3\frac{\tau-\xi}{\tau-\zeta}\right] d\zeta. \quad (30)$$

One may obtain the following inequalities from Eq. (30) when R is a monotonic increasing function of time:

$$-\left(\frac{D}{\pi}\right)^{\frac{1}{2}} \frac{1}{R^3(\tau)} \int_0^\tau \frac{\theta^0(0,\zeta)}{(\tau-\zeta)^{\frac{1}{2}}} d\zeta \le T'(0,\tau) \le -\frac{4}{3}\left(\frac{D}{\pi}\right)^{\frac{1}{2}} \times \int_0^\tau \frac{\theta^0(0,\zeta)}{R^3(\zeta)(\tau-\zeta)^{\frac{1}{2}}} d\zeta. \quad (31)$$

When R is a monotonic decreasing function, the sense of the inequalities should be reversed. It may be noted in Eq. (31) that

$$-\left(\frac{D}{\pi}\right)^{\frac{1}{2}} \int_0^\tau \frac{\theta^0(0,\zeta)}{(\tau-\zeta)^{\frac{1}{2}}} d\zeta = D\int_0^\tau F(\zeta)d\zeta,$$

where $F(\zeta)$ is defined by the boundary condition of Eq. (9). When R is a monotonic increasing function, Eq. (31) thus gives

$$\frac{D}{R^3}\int_0^\tau F(\zeta)d\zeta \le T'(0,\tau) \le \frac{4D}{3R_0^3}\int_0^\tau F(\zeta)d\zeta, \quad (32)$$

where $R_0 = R(0)$. The upper bound in Eq. (32) represents, of course, a poorer estimate than that given in Eq. (31).

V. CONCLUSION

The approximations developed here have been applied by the authors to the problem of the growth of a vapor bubble in a superheated liquid. For this specific problem, one can examine in detail the validity of the assumption of the thin "thermal boundary layer" which has been justified previously only in general physical terms. Such an examination of the predicted temperature field shows that the zero-order approximation, as given by Eq. (17) or (17′), is sufficient. Therefore, an explicit expression for the first-order temperature correction at any point in the liquid has not been given here, although it may be found from Eq. (26). The first-order temperature correction at the boundary $r=R(t)$ is given by a fairly simple expression, and the given bounds upon it provide a convenient estimate of the rapidity of the convergence of the approximation theory.

The approximation procedure presented here is not limited to heat diffusion across a spherical vapor bubble in a liquid. The theory applies without alteration to diffusion across any spherical boundary with radial motion in a fluid, provided the thin "thermal boundary layer" approximation is valid. It is to be emphasized that effects of any translational motion of the spherical boundary have not been considered. For the case of the vapor bubble in a liquid, the solution is therefore applicable only over time intervals so short that no significant translation can take place.

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 1 JANUARY, 1952

The Temperature Dependence of Viscosity for Water and Mercury*

PRESTON M. KAMPMEYER†
Film Research and Development Division, Olin Industries, Inc., New Haven, Connecticut
(Received August 8, 1951)

Smoothed values for the viscosities of water and mercury, as reported in the literature, show surprisingly large percentage deviations from the experimental data. Because water is frequently used as the calibrating fluid in viscometry, it was of interest to apply expressions, suggested by the theory of Eyring, for smoothing experimental data. It was of further interest to apply the Arrhenius expression for the variation of viscosity with temperature to the data for both water and mercury and to compare the closeness of fit with that obtained by the use of the aforementioned expressions. As judged by the Gaussian criterion of fit, an expression suggested by the Eyring theory provides smoothed values for the viscosities of both water and mercury which are believed to be more reliable than those in common use. This is especially true for the viscosity of water, for which relative viscosity data were first smoothed and then pinned to the recent value of 10.019±0.003 millipoises at 20°C as reported by Swindells and Coe.

From other empirical equations relating densities and energies of vaporization to temperature, equations are derived for expressing the activation energies for viscous flow as a function of temperature.

I. INTRODUCTION

SMOOTHED values for the viscosities of water and mercury, as reported in the literature, show surprisingly large percentage deviations from the experimental data.[1–3] Since it is now widely recognized that viscosity data can be fairly well represented by a plot of the logarithm of viscosity *vs* the reciprocal of the absolute temperature,[4,5] it was of interest, and of importance in connection with work reported elsewhere,[6] to subject good viscosity data to curve fitting techniques, including least squaring procedures, utilizing the Arrhenius equation in expanded form,

$$\log\eta = a+bx+cx^2+dx^3+\cdots, \tag{1}$$

where $x=10^3/T$; and equations of the form,

$$\log(\eta/T^{\frac{1}{2}}) = aT^{-4}+bT^{-3}+\cdots+p+qT+rT^2+\cdots, \tag{2}$$

suggested by the theory of Eyring.

In making the choice of the most suitable equation to represent the data, and in limiting the number of constants in the equation chosen, the Gaussian criterion of fit[7] was the deciding factor.

In the theory of Eyring,[8] treating viscosity as a rate process, the dependence of viscosity upon temperature is expressed by the equation,

$$\eta = (qM^{1/2}T^{3/2}/V^{2/3}\Delta E_{\text{vap}})e^{E/RT} \tag{3}$$

or, by the equation,

$$\eta = (qd^{2/3}T^{3/2}/M^{1/6}\Delta E_{\text{vap}})e^{E/RT}, \tag{4}$$

where $q=N^{1/6}cR(2\pi k)^{1/2}$; c is a number characteristic of the type of molecular packing in the liquid and may be temperature dependent, M = molecular weight, V = molar volume, d = density, E = molar energy of vaporization; and N, R, and k have their usual meanings.

If Eq. 4 is written in the form,

$$\ln(\eta/T^{3/2}) = \ln\phi(T)+\ln q+E/RT,$$

where

$$\phi(T) = (d^{2/3}/M^{1/6}\Delta E_{\text{vap}}), \tag{5}$$

one may write,

$$[\ln(\eta/T^{3/2})-\ln\phi(T)-\ln q]RT = E(T), \tag{6}$$

$E(T)$ signifying that the activation energy, E, may be some function of temperature.

For water and mercury, it was found that equations of the form,

$$\phi(T) = a+bT+cT^2 \tag{7}$$

and

$$E(T) = aT^{-1}+b+cT+\cdots \tag{8}$$

serve adequately in fitting the data. One notes that, granted the validity of the Eyring theory as expressed by Eq. (4), Eq. (8) expresses the activation energy as a function of temperature.

II. THE VISCOSITY OF WATER

In considering the viscosity of water, two separate sets of data were used: (1) the classic data of Bingham and White,[9] who determined absolute viscosities at 5° intervals from 0° to 95°C, and (2) the relative viscosi-

* Taken from Part B of a dissertation in chemistry submitted to the Faculty of the Graduate School of the University of Pittsburgh in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

† Present address: Olin Industries, Inc., P. O. Drawer 906, New Haven, Conn.

[1] *International Critical Tables* (McGraw-Hill Book Company, Inc., New York, 1929).

[2] N. E. Dorsey, *The Properties of Ordinary Water Substance*, Reinhold Publishing Company, New York, (1940).

[3] E. C. Bingham and R. F. Jackson, Bull. Bur. Standards, **14**, 74 (1917).

[4] Glasstone, Laidler, and Eyring, *The Theory of Rate Processes* (McGraw-Hill Book Company, Inc., New York, 1941), p. 484.

[5] Lewis, Squires, and Broughton, *Industrial Chemistry of Colloidal and Amorphous Materials* (Macmillan Company, New York, 1944), p. 29.

[6] Kampmeyer, Mason, and Robinson, *Viscosities of Aqueous Solutions of Amino Acids at 25° and 35°C* (to be published).

[7] Archie G. Worthing and Joseph Geffner, *Treatment of Experimental Data* (John Wiley and Sons, Inc., New York, 1946), p. 261.

[8] Reference 4, Eq. 35, p. 486.

[9] E. C. Bingham and G. F. White, Z. phvsik. Chem. **83**, 641 (1913).

TABLE I. The viscosity of water (data of Bingham and White): values for the constants in Eqs. 1 and 9 (T = °C + 273.16)

Equation	Constants			
	a	b	c	d
1[a]	−5.485471	5.041873	−1.551439	0.1849577
9[b]	−5.261387	4.586889	−1.3229745	0.1589945

(a) Viscosity in poise × 10^3
(b) Viscosity in poise × 10^7

ties as reported by recent workers[10–12] and the value of 10.019±0.003 millipoises for the absolute viscosity of water at 20°C now used at the Bureau of Standards as recently reported by Swindells and Coe.[13]

Bingham and Jackson[3] least squared Bingham and White's data[9] to fit the equation,

$$1/\eta = A\{(t-B)+[C+(t-B)^2]^{\frac{1}{2}}\}-D,$$

chosen because it minimized the differences between smoothed and observed values; these authors also included in the data chosen for least squaring much of the viscosity data obtained prior to 1913.

In fitting the data of Bingham and White to Eq. (1) with three constants, a negative value was obtained for b, and the differences $y_0 - y_c$ showed a trend. With an added constant, a positive value was obtained for b, the differences $y_0 - y_c$ exhibited random scattering, and the quantities $(y_0 - y_c)/y_0$ were all less than one percent; application of the Gaussian criterion of fit showed that the four-constant equation was justified.

In fitting the data of Bingham and White to Eq. (2), it was found that the best fit was obtained when the equation took the form,

$$\log(\eta/T^{\frac{1}{2}}) = a + bx + cx^2 + dx^3, \qquad (9)$$

where again $x = 10^3/T$. This was also found to be the case when treating the data for relative viscosities.

Some results of these smoothing operations are summarized in Table I. In Table II are tabulated the observed values of Bingham and White, smoothed values as obtained by Bingham and Jackson, and smoothed values as obtained by use of Eq. (9).

In recent years several apparently reliable values have been reported for the relative viscosities of water (relative viscosity = viscosity of water at t°C relative to the viscosity of water at 20°C), for temperatures ranging from 25 to 60°C.[10–12] These relative values, taken in conjunction with the very recent value of 10.019±0.003 millipoises for the absolute viscosity of water at 20°C as reported by Swindells and Coe,[13] afford the opportunity for obtaining very accurate values for absolute viscosities of water for a wide range of temperatures.

In treating these data, it was found that the best form of Eq. (2) was Eq. (9), after making the substitution of η_r for η. *A priori* weights were assigned to the various values of η_r on the following basis.

The values of η_r reported for all temperatures except 20 and 25° seemed of equal worth and were, accordingly, initially assigned the weight 1. Because the same value (0.8885) for η_r at 25° was obtained by each of four investigators, it was assigned the weight 4. The value of 1 for η_r at 20° is a defined value and was assigned the weight 10.

Assigning *a priori* weights to the various values of η_r, as explained above, it was found that the value of η_r at 20°, calculated from the equation obtained, deviated from the defined value of 1.0000 by 0.3 percent, indicating either that Eq. (9) was inadequate for representing the data, or that one or more of the values for the relative viscosity was seriously in error. Inspection indicated the desirability of eliminating from the least squaring operation the value for the relative viscosity of water at 30° as reported by Geddes. When this was done, the value for η_r calculated from the least squares equation did not differ significantly from the defined value of 1.0000.

The results from this latter least squaring operation are shown in Table III. Using the relative viscosities so obtained, together with the value of 10.019 millipoises as reported by Swindells and Coe, values for the viscosities of water over the temperature range 0–100° were calculated and are compared, in Table III, with the experimental values of Bingham and White. The values for the viscosities of water shown in column 3 of Table III are believed to be as reliable as any hitherto reported, especially for the range 20–60°.

Several items seem worthy of mention in connection with the curve fitting process as applied to the viscosity data for water. In all attempts to fit the data of

TABLE II. Observed and smoothed values for the viscosity of water.

°C	η_0[a]	$(\eta_0-\eta_c)100/\eta_0$[b]	$(\eta_0-\eta_c)100/\eta_c$[c]
0	17.960	0.217	−0.189
5	15.241	0.348	−0.230
10	13.002	−0.577	0.023
20	10.054	0.040	−0.536
25	8.940	0.034	0.089
30	7.991	−0.200	−0.088
35	7.223	−0.028	0.125
40	6.557	−0.046	0.107
45	5.984	−0.067	0.050
50	5.491	−0.055	0.018
55	5.073	0.177	0.217
60	4.728	0.846	0.888
65	4.362	0.160	0.183
70	4.069	0.197	0.245
75	3.794	−0.132	0
80	3.558	−0.196	0.036
85	3.337	−0.539	−0.120
90	3.133	−1.021	−0.383
95	2.983	−0.368	0.570

[a] Data of Bingham and White, viscosity in millipoises.
[b] The η_c are the smoothed values of Bingham and Jackson.
[c] Here the η_c are smoothed values obtained from Eq. 9 whose constants are given in Table I.

[10] J. R. Coe, Jr., and T. B. Goeffry, J. Appl. Phys. **15**, 625 (1944).
[11] J. A. Geddes, J. Am. Chem. Soc. **55**, 4832 (1933).
[12] Reference 10, footnote "*h*" in Table I (work of Swindells at the Bur. Standards).
[13] J. F. Swindells and J. R. Coe, Jr., Rheology Bull. **19**, No. 2, p. 3 (Fall, 1950).

Bingham and White to various forms of Eqs. (1) and (2) by the method of successive approximation, the point for 60° was always a "wild" point in all deviation plots. To a lesser extent, and depending on the exact form of Eqs. (1) and (2) used, points for 0, 5, 90, and 95° were also "wild" points in deviation plots. The experimental value for the viscosity of water at 0°C reported by Bingham and White, *viz.*, 17.960 millipoises, is higher than any of the five other values reported by previous workers.[3]

Equation (9) is a successful extrapolation formula in the sense that, having determined the arbitrary constants with the use of data for the temperature range 20–60°, it successfully predicts viscosities in the range 60–100° and in the range 0–20°.

III. THE VISCOSITY OF MERCURY

Interest in the viscosity of mercury arose from several considerations. Aside from the intrinsic worth of obtaining a reliable value for its coefficient of viscosity over a wide range of temperature, it was of interest to test Eqs. (1) and (2) with available data for a simple liquid. Further interest arose from the possibility of comparing its activation energy for viscous flow, and the change of this property with temperature, with the activation energy for the viscous flow of water.

The data chosen for the smoothing operation are those of Plüss;[14] these were chosen because they seemed to be as reliable as any other data reported for the viscosity of mercury and extend over the widest range of temperature investigated by any single worker.

Because mercury is probably one of the simplest liquids known, and because Plüss' data extend over such a large temperature range, a good test is afforded

TABLE III. The relative viscosities of water calculated from the least squares equation: $\log(\eta_r \times 10^5/T^4) = -3.1624338 + 2.6056821/T - 0.70113982/T^2 + 0.09402149/T^3$; the absolute viscosities of water based on the value of 10.019 millipoises for the viscosity of water at 20°C (T = °C + 273.16).

°C	η_r	η_c	(Bingham and White) η_0	(Bingham and White) $(\eta_c - \eta_0)100/\eta_c$
0	1.7668	17.702	17.960	−1.457
5	1.5079	15.108	15.241	−0.880
10	1.3014	13.039	13.002	0.284
15	1.1353	11.374	11.373	0.009
20	1.0000	(10.019)	10.054	−0.349
25	0.88858	8.9027	8.940	−0.415
30	0.79575	7.9726	7.991	−0.226
35	0.71767	7.1903	7.223	−0.459
40	0.65139	6.5263	6.557	−0.475
45	0.59603	5.9716	5.984	−0.201
50	0.54571	5.4675	5.491	−0.421
55	0.50319	5.0415	5.073	−0.615
60	0.46599	4.6688	4.728	−1.264
65	0.43325	4.3407	4.362	−0.484
70	0.40426	4.0503	4.069	−0.469
75	0.37846	3.7918	3.794	−0.053
80	0.35536	3.5604	3.558	0.056
85	0.33461	3.3524	3.337	0.448
90	0.31587	3.1647	3.133	1.010
95	0.29888	2.9945	2.983	0.368
100	0.28395	2.8395	...	...

[14] M. Plüss, Z. anorg. allgem. Chem. **83**, 1 (1915).

TABLE IV. Representation of Plüss' data for the viscosity of mercury (in millipoise) by the least squares equations: (1) $\log\eta = a + bT^{-1}$; (2) $\log\eta = a + bT^{-1} + cT^{-2}$; (3) $\log(\eta \times 10^4/T^4) = a + bT^{-1} + cT + dT^2$; (4) $\log(\eta \times 10^4/T^4) = a + bT^{-1} + cT^{-2} + dT^{-3}$.

Equation	Arbitrary constants a	b	c	d	Max. dev., %
1	0.69811	149.31	0	0	6.68
2	0.68893	183.81	10880	0	2.83
3	1.87744	123.87	−0.0034038	2.17857×10^{-6}	2.51
4	−0.25969	830.83	−129900.5	10,729,950	2.45

the simple Arrhenius equation, and extensions of this equation such as Eq. (1). A summary of some of the results obtained by the least squaring of Plüss' data to fit Eq. (1) is contained in Table IV. It seems worthwhile to point out that the curve obtained by fitting the data to the Arrhenius equation is markedly skewed, and that the percentage difference between observed and smoothed value is as large as 6.68 percent.

As in the case of water, the best form of Eq. (2) for use in fitting the data was found to be Eq. (9). The results of some of the smoothing operations are summarized in Table IV. Plüss' data are given in Table V together with the smoothed values obtained by use of Eq. (9).

IV. ACTIVATION ENERGY FOR VISCOUS FLOW AND ITS CHANGE WITH TEMPERATURE

An activation energy for the viscous flow of a liquid might be expected to be related to the nature and strength of the interaction forces between, and to the size and geometrical arrangement of, the molecules of which it is composed. It seems reasonable, therefore, to expect that activation energy will be a function of temperature.

It was previously pointed out that, granted the validity of Eq. (4), Eq. (8) expresses the activation energy as a function of temperature. It remains, therefore, to evaluate the parameter $\phi(T)$ defined by Eq. (5). As previously stated, $\phi(T)$ is adequately represented as a function of temperature by Eq. (7).

In applying Eqs. (5) and (7) to the data for mercury, Menzies' equation for the vapor pressure of mercury was used,[15] together with the Clausius-Claperyron equation (in differential form) to obtain molar heats of vaporization, ΔH. In applying the same equations to the data for water, molar heats of vaporization were used as reported by Keyes and Smith.[16] In both cases, using values for densities reported by the Smithsonian Institution and the relation $\Delta E = \Delta H - RT$, Eq. (7) was successful in representing the parameter $\phi(T)$.

The explicit forms of Eq. (7) for mercury and water are as follows:

For mercury,

$$\phi(T) \times 10^5 = 16.2737 + 31.375 \times 10^{-4} T + 1.92188 \times 10^{-6} T^2 \qquad (T = °C + 273.2);$$

[15] A. W. C. Menzies, J. Am. Chem. Soc. **41**, 1783 (1919).
[16] F. G. Keyes and L. B. Smith, Mech. Eng. **53**, 132 (1931).

TABLE V. Plüss' data for the viscosity of mercury: smoothed values obtained by the use of Eq. 9. (T = °C + 273.2).

°C	η_0	η_c[a]	$(\eta_0-\eta_c)100/\eta_0$
0	16.63	16.67	−0.24
0	16.59	16.67	−0.48
16.5	15.72	15.66	0.38
18.2	15.82	15.56	1.64
18.6	15.80	15.54	1.65
19.5	15.12	15.49	−2.45
20.0	15.47	15.47	0
33.8	14.82	14.78	0.21
34.2	14.70	14.75	−0.34
62.9	13.603	13.635	−0.23
79.4	12.99	13.11	−0.92
97.5	12.70	12.60	0.78
98.4	12.56	12.58	−0.16
136.8	11.78	11.72	0.51
137.5	11.78	11.70	0.68
138.2	11.47	11.69	−1.92
176.9	11.19	11.03	1.43
180.0	10.845	10.986	−1.30
180.0	11.13	10.986	1.26
191.0	10.95	10.83	1.09
193.0	10.84	10.80	0.37
194.5	10.58	10.78	−1.89
228.0	10.35	10.38	−0.29
262.0	10.05	10.04	0.01
299.0	9.751	9.753	−0.02

[a] From Eq. 4 of Table IV.

maximal difference between smoothed and "observed" value is 0.01 percent and occurs at 320°C.

For water,

$$\phi(T)\times 10^7 = 2.02082 + 2.9049\times 10^{-3}T \qquad (T = °C + 273.16);$$

maximal difference between smoothed and "observed" value is 0.07 percent and occurs at 40°C.

It has been stated that c, a number characteristic of the type of molecular packing in the liquid, may be temperature dependent. For cubical packing c has the value 2;[17] if the assumption is now made that c retains this value for both mercury and water in the temperature range for which viscosity data are to be considered, then q may be evaluated and there remains only the task of evaluating the constants in Eq. (8) in order to obtain the activation energy as a function of temperature.

For water, it was found that the activation energy for viscous flow may be represented by the equation,

$$E_{act} = 5140.4 - 2.5810t + 0.06063t^2 - 2.4138\times 10^{-4}t^3, \qquad (t = °C).$$

Activation energies at 0, 25, 35, and 100°, calculated from this equation, are, respectively, 5140.4, 5110.0, 5114.0, and 5247.0 cal per mole, and the activation energy is minimal at 25°. The decrease in activation energy with increasing temperature, up to 25°, may be a reflection of the extent to which water becomes depolymerized with increasing temperature. No explanation is at hand for the increase in activation energy as the temperature increases beyond 25°.

The activation energy for the viscous flow of mercury obtained, as in the case of water, from the expression,

$$E_{act} = 2.303RT(\log(\eta/T^{\frac{1}{2}}) - \log\phi(T) - \log q)$$

may be represented by the equation,

$$E_{act} = 1817.5 + 1.397T - 0.00284722T^2.$$

The activation energy in this case decreases with increasing temperature for viscosity data in the range 0–300°C.

In the case of mercury, it was of interest to define a new parameter,

$$\theta(T) = (d^{2/3}/M^{1/6}\Delta H_{vap}).$$

In this case, Eq. (7) was strikingly successful in representing the parameter $\theta(T)$, the explicit form of the equation being—

$$\theta(T)\times 10^5 = 16.2263 + 12.9666\times 10^{-4}T + 3.960055\times 10^{-7}T^2, \qquad (T = °C + 273.2);$$

maximal difference between smoothed and "observed" value is 0.004 percent and occurs at 320°C. With this new parameter, it becomes possible to express the temperature dependence of the viscosity of mercury by an equation containing but one arbitrary constant. The fit is almost as good as that obtained by the use of the two constant Arrhenius equation for the same range of temperature, the maximum deviation being 7.69 percent *vs* 6.68 percent for the Arrhenius equation, but the scattering of the deviations is more random. The equation is

$$\log(\eta\times 10^8/T^{\frac{1}{2}}) = \log\theta(T)\times 10^5 = (373.03/T).$$

Omitting the viscosity value for 299° in the curve fitting process, the fit obtained by the use of the above equation is considerably better than the fit obtained by the use of the Arrhenius equation as judged by the Gaussian criterion. This use of a equation containing but one arbitrary constant to fit viscosity data extending over a temperature range of greater than 250°, giving a maximum deviation of less than 5 percent, is rather remarkable.

VI. ACKNOWLEDGMENTS

The author wishes to express his appreciation to Miss Phyllis Alspach for assistance with some of the more tedious computations. Appreciation is also expressed to members of the Department of Chemistry at the University of Pittsburgh and to Dr. John R. Bowman of the Mellon Institute, Pittsburgh, for valuable discussions and constant encouragement.

[17] Reference 4, p. 479.

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 1 JANUARY, 1952

Instantaneous Power Spectra

CHESTER H. PAGE
National Bureau of Standards, Washington, D. C.
(Received August 16, 1951)

The intuitive concept of a changing spectrum is discussed. The *instantaneous power spectrum* is defined mathematically and used to make the intuitive concepts more precise. It depends upon the past history of a signal, but not upon the future.

Integration of the instantaneous power spectrum over time yields the conventional energy spectrum. The instantaneous power spectrum of a random function may be averaged over the ensemble of functions, with a resulting stochastic average instantaneous power spectrum that is equal to the conventional time average power spectrum of a stochastic process.

THE ordinary concept of the frequency spectrum of a signal is often confusing in practice. The spectrum is expressed in terms of the Fourier transform of the signal; the Fourier transform is determined by the complete history of the signal from $t=-\infty$ to $t=+\infty$. In practice, we are interested in what is happening *now*, and this is independent of what the signal may do in the future. It is a familiar fact that in the Fourier solution of a circuit problem, the future behavior of the applied signal does not affect the result, even though it apparently enters the calculation via the transform.

The (energy) spectrum of a signal is defined as the absolute square of the Fourier transform of the signal. Thus a signal $G(t)$ has the Fourier transform

$$g(f)=\int_{-\infty}^{\infty} e^{-2\pi ift}G(t)dt \qquad (1)$$

if the integral converges. Plancherel's theorem states that

$$\int_{-\infty}^{\infty}|g(f)|^2df=\int_{-\infty}^{\infty}|G(t)|^2dt. \qquad (2)$$

Since $G^2(t)$ is the instantaneous power (into unit resistance), (2) represents the total energy of the signal, and $|g|^2$ is interpreted as the energy distribution in frequency, or the energy spectrum. Thus the spectrum depends upon the entire history of the signal, by virtue of the infinite limits in (1). This property of the spectrum is responsible for many apparent paradoxes and much confusion in description.

For example, let a man and woman sing alternate verses of a song. It is a common intuitive reaction to say that the spectrum keeps changing. We somehow feel that a sample of sufficient duration defines a spectrum, yet the mathematical definition of spectrum disallows this. The paradox becomes more apparent when we decrease the time intervals for spectral comparison and ask how the spectrum changes in successive milliseconds. This is "obviously" an absurd query. The paradox is somewhat resolved by a sharper analysis. By the spectrum of the male verses we really mean the power spectrum of this man's singing, or mathematically the power spectrum of the stochastic process of which the singing of one verse is a statistically large sample. This standard concept of power spectrum will be discussed later. The role of duration is now clarified; one verse is statistically sufficient to give a reasonably precise determination of the process and its spectrum, one syllable is not a large enough sample.

The paradox has not, however, been completely vanquished. There is certainly nothing statistical about a pure sine wave. Let us play a simple melody on a very crude electric organ, or more ideally, change the frequency of an audio oscillator. We describe both these situations in terms of frequency changes in time. Again, each tone must be of sufficient duration to define a frequency. But how long does it take to establish a frequency? Indeed, what do we mean by a frequency, since any signal of finite duration has a spectrum that is not infinitely sharp? A unique frequency exists only for a sine wave of infinite duration.

There must be some description of signals that asymptotically agrees with the above intuitive properties. That is, a sine wave switched on at $t=0$ has some sort of spectral distribution that can be represented more and more closely by a single frequency as time goes on. Similarly, a stochastic process can be spectrally described in some manner that progressively agrees more closely with the power spectrum of the process. We seek a description in terms of an "instantaneous power spectrum" with the above properties.

Consider the energy of a signal to be distributed in time and frequency. Denote the energy density in this $t-f$ plane by $\rho(t, f)$. The total energy expended up to time T will be

$$\int_{-\infty}^{T}\int_{-\infty}^{\infty}\rho(t, f)dfdt.$$

The rate of increase of total energy is the instantaneous power, and is simply

$$\int_{-\infty}^{\infty}\rho(T, f)df$$

by differentiating the energy. Thus at any time T, the

distribution $\rho(T, f)$ can be considered as the instantaneous power spectrum. Suppose the signal $G(t)$ to be switched off at $t=T$. This defines an auxiliary signal

$$G_a(t)=\begin{cases}G(t) & t<T\\ 0 & t>T,\end{cases}$$

which is identical with $G(t)$ up till time T, and therefore up till time T has delivered the same energy as $G(t)$. Since G_a is defined for all time, we can compute its Fourier transform and its energy spectrum, and the latter should agree with the time integrated power spectrum of $G(t)$. That is, we require $\rho(t, f)$ to satisfy the condition

$$\int_{-\infty}^{T}\rho(t, f)dt=|g_a(f)|^2,$$

where $g_a(f)$ is the transform of $G_a(t)$. Since this condition must be satisfied for *all* T, it is convenient to define the *running transform* of $G(t)$ as the transform of a continually changing auxiliary function:

$$G_t(x)\equiv\begin{cases}G(x) & x<t\\ 0 & x>t\end{cases} \tag{3}$$

$$g_t(f)\equiv\int_{-\infty}^{\infty}G_t(x)e^{-2\pi ifx}dx=\int_{-\infty}^{t}G(x)e^{-2\pi ifx}dx. \tag{4}$$

The instantaneous power spectrum must now satisfy

$$\int_{-\infty}^{t}\rho(x, f)dx=|g_t(f)|^2 \tag{5}$$

and this equation is *sufficient* to determine ρ. Differentiation with respect to t yields

$$\rho(t, f)=(\partial/\partial t)|g_t(f)|^2. \tag{6}$$

The expression for $\rho(t, f)$ can profitably be written in other forms. First, substitute (4) into (6), so that

$$\rho(t, f)=(\partial/\partial t)\left\{\int_{-\infty}^{t}G(x)e^{-2\pi ifx}dx\int_{-\infty}^{t}G(y)e^{+2\pi ify}dy\right\}$$

$$=G(t)e^{-2\pi ift}\int_{-\infty}^{t}G(y)e^{+2\pi ify}dy$$

$$+G(t)e^{2\pi ift}\int_{-\infty}^{t}G(x)e^{-2\pi ifx}dx \tag{6a}$$

$$=2G(t)Re:e^{2\pi ift}g_t(f). \tag{7}$$

This is usually the most convenient form for calculation of $\rho(t, f)$.

The intermediate result (6a) can be used as follows:

$$\int_{-\infty}^{\infty}\rho(t, f)df=G(t)\left\{\int_{-\infty}^{\infty}\int_{-\infty}^{t}G(y)e^{2\pi if(y-t)}dydf\right.$$

$$\left.+\int_{-\infty}^{\infty}\int_{-\infty}^{t}G(x)e^{-2\pi if(x-t)}dxdf\right\}$$

$$=G(t)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\{G_t(x)e^{2\pi if(x-t)}+G_t(x)e^{-2\pi if(x-t)}\}dxdf$$

$$=2G(t)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}G_t(x)\cos 2\pi f(x-t)dxdf.$$

By Fourier's theorem, the double integral is

$$\tfrac{1}{2}[G_t(t-0)+G_t(t+0)]=\tfrac{1}{2}G(t)$$

by virtue of the discontinuity in $G_t(x)$ by definition (3). Therefore

$$\int_{-\infty}^{\infty}\rho(t, f)df=G^2(t)=\text{instantaneous power},$$

as expected.

[Equation (6a) can also be written as

$$\rho(t, f)=2G(t)\int_{-\infty}^{t}G(x)\cos 2\pi f(x-t)dx$$

and a change of variable yields

$$\rho(t, f)=2G(t)\int_{0}^{\infty}G(t-\tau)\cos 2\pi f\tau d\tau$$

$$=2\int_{0}^{\infty}G(t)G(t-\tau)\cos 2\pi f\tau d\tau \tag{8}$$

where $G(t)$ has been placed inside the integral purely for later convenience.]

Consider the spectrum generated by a step-wave, i.e., direct current switched on at $t=0$. We find directly

$$g_t(f)=\int_{0}^{t}e^{-2\pi ifx}dx=(e^{-2\pi ift}-1)/(-2\pi if)$$

$$Re:e^{2\pi ift}g_t(f)=Re(1-e^{2\pi ift})/(-2\pi if)=(\sin 2\pi ft/2\pi f)$$

and

$$\rho(t, f)=2(\sin 2\pi ft/2\pi f). \tag{9}$$

This spectrum is of the form $\sin x/x$ with its familiar oscillations. For any t, ρ has its greatest maximum at $f=0$, and this value ($\rho_m=2t$) grows indefinitely. As time goes on, the distribution becomes more concentrated around $f=0$; the peaks of the instantaneous spectrum move in toward $f=0$ maintaining constant

area under each "hump" of the curve. The area

$$\int_{-\infty}^{\infty} \rho df = 1$$

for all t, and the spectrum approaches a δ-function as $t \to \infty$. Thus as t increases, the spectrum is more and more precisely described by stating that the signal is an oscillation of zero frequency.

The stepped sine wave

$$G(t) = \begin{cases} 0 & t<0 \\ \sin 2\pi f_0 t, & t>0 \end{cases}$$

exhibits a similar behavior. Its spectrum is found by (7) to be

$$\rho(f, t) = \frac{1}{\pi} f_0 \sin 2\pi f_0 t (\cos 2\pi f_0 t - \cos 2\pi f t)/(f^2 - f_0^2)$$

$$= \frac{2}{\pi} f_0 \sin 2\pi f_0 t \frac{\sin \pi (f+f_0)t}{f+f_0} \frac{\sin \pi (f-f_0)t}{f-f_0}.$$

As $t \to \infty$, the spectrum becomes a δ-function distribution at $f=f_0$ and $f=-f_0$. The appearance of negative frequencies is not conceptually troublesome because the spectrum is always an even function of f. Thus negative frequencies can be ignored in the interpretation.

Negative energy densities do occur, as is obvious in (9). These seem physically unreal at first glance, but must be present. For if either of these stepped signals is turned off when t is small, the total spectrum contains high frequencies. If turned off later, the high frequency content is less, but the high frequency content must have been present originally. This means that as time goes on, the "virtual" sidebands must be modified and some of the energy recalled. Negative values of the density ρ occur only in partial compensation of earlier positive ρ; the total energy $\int_{-\infty}^{t} \rho dt$ is always positive at all frequencies, by Eq. (5).

Another interesting signal is the isosceles triangle, such as

$$G(t) = \begin{cases} 0 & t<-1 \\ 1+t, & -1<t<0 \\ 1-t, & 0<t<1 \\ 0, & 1<t. \end{cases}$$

The instantaneous power spectrum is readily found

$$\rho(t, f) = (1+t)\{1-\cos 2\pi f(1+t)\}/2\pi^2 f^2 \quad (10a)$$

for $-1<t<0$

$$\rho(t,f) = (1-t)\{2\cos 2\pi f t - 1 - \cos 2\pi f(1+t)\}/2\pi^2 f^2 \quad (10b)$$

for $0<t<1$ and $\rho(t, f)=0$ elsewhere. During the rise of the signal, the energy density is everywhere positive by (10a). During the latter half of the signal, there are regions of negative density. At all times, the density is maximum at zero frequency, and this maximum is greatest shortly after $t=0$. The direct component of the signal is still increasing when the signal peak is passed. A relief plot of ρ is shown in Fig. 1.

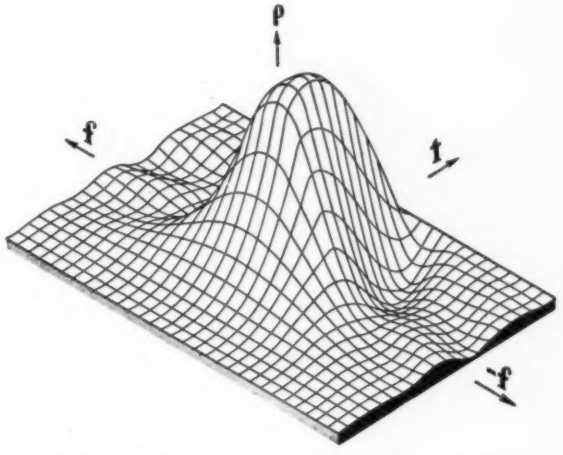


FIG. 1. Energy distribution of a triangular pulse.

EXTENSION TO REACTIVE CIRCUITS

The definition of power spectrum in terms of pure (unit) resistance load was made for simplicity. Consider the voltage $E(t)$ applied across a circuit containing reactance, with the resulting current $I(t)$. The instantaneous power delivered by the source is EI, so the total energy delivered is

$$\int_{-\infty}^{\infty} E(t)I(t)dt = \int_{-\infty}^{\infty} e(f)\bar{\imath}(f)df$$

$$= \int_{-\infty}^{\infty} \bar{e}(f)i(f)df \quad (11)$$

$$= \int_{-\infty}^{\infty} Re:(e\bar{\imath})df,$$

where $e(f)$ and $i(f)$ are the Fourier transforms of $E(t)$ and $I(t)$, with $\bar{e}$ the complex conjugate of e. The equation follows from Plancherel's theorem, and the energy spectrum can be taken as any of the three integrands on the right. The complex forms yield an imaginary component of energy density. This component is an odd function of frequency, so its net power is zero. Inclusion of this term does not apparently have any usefulness, so we take the energy density in the third form: real part of $e\bar{\imath}$.

The energy delivered by the source up until time t is

$$E = \int_{-\infty}^{t} E(x)I(x)dx = \int_{-\infty}^{\infty} E_t(x)I_t(x)dx \quad (12)$$

where $I_t(x)$ is the current resulting from the auxiliary voltage $E_t(x)$ which vanishes for $x>t$. Note that $I_t(x)$ does not vanish identically for $x>t$, and note also that

any arbitrary definition of $I_t(x)$ for $x>t$ would not affect Eq. (12). By Plancherel's theorem, the right-hand side of (12) is equal to

$$E=\int_{-\infty}^{\infty} Re:e_t(f)\bar{\imath}_t(f)df$$

so that the cumulative energy spectrum can be taken as $Re:e_t(f)\bar{\imath}_t(f)$. Its rate of increase is the instantaneous power spectrum

$$\rho(t, f)=(\partial/\partial t)Re:e_t(f)\bar{\imath}_t(f). \tag{13}$$

The running transform of the current $i_t(f)$ can be found in terms of the impedance of the circuit, or its reciprocal, the admittance:

$$i_t(f)=e_t(f)/z(f)=e_t(f)A(f). \tag{14}$$

This yields

$$\rho(t, f)=(\partial/\partial t)Re:|e_t(f)|^2A(f)=C(f)(\partial/\partial t)|e_t(f)|^2 \tag{15}$$

using $A(f)=C(f)+iS(f)$. The result (15) shows that $\rho(t, f)$ has the original form (6) multiplied by the conductance. *This is the spectrum for the power delivered by the source, not that dissipated by the resistance.* These two are not in continual balance because energy is stored in the reactance. The corresponding total energy spectra for $t=\infty$ must, however, be equal.

RANDOM FUNCTIONS

When the driving force $G(t)$ is a random function, a continuing process, the total energy becomes infinite with time. The Fourier transform and the conventional energy spectrum do not exist. An average energy spectrum per second, or power spectrum, does exist and its customary definition can be expressed in terms of our running transform:[1]

$$P(f)=\lim_{T\to\infty}(1/T)|g_T(f)|^2. \tag{16}$$

The usual procedure for finding $P(f)$ is to use the Wiener-Khintchine theorem:[1]

$$P(f)=2\int_0^{\infty}\Psi(\tau)\cos 2\pi f\tau d\tau$$

where

$$\Psi(\tau)\equiv\lim_{T\to\infty}(1/T)\int_0^T G(t)G(t-\tau)dt$$

is, except for normalization, the auto-correlation function of the process $G(t)$. Thus the standard form for the power spectrum of a stochastic process is

$$P(f)=2\int_0^{\infty}\langle G(t)G(t-\tau)\rangle_{Av^{(t)}}\cos 2\pi f\tau d\tau \tag{17}$$

with $\langle G(t)G(t-\tau)\rangle_{Av^{(t)}}$ indicating the *time average* of the product.

Now for a stationary ergodic process, *time averages* and *stochastic averages* over the ensemble should be equivalent. Hence we expect that our instantaneous power spectrum formula, applied to random functions, will yield a result corresponding to (17) after stochastic averaging. Indeed, the stochastic average of (8) is

$$\langle\rho(t, f)\rangle_{Av^{(s)}}=2\int_0^{\infty}\langle G(t)G(t-\tau)\rangle_{Av^{(s)}}\cos 2\pi f\tau d\tau$$

which is the power spectrum $P(f)$ if G is an ergodic process.

Returning to Eq. (8), if $G(t)$ is a function switched on at $t=0$, the integrand vanishes for $\tau>t$, and

$$\rho(t, f)=2\int_0^t G(t)G(t-\tau)\cos 2\pi f\tau d\tau. \tag{18}$$

Hence for a random function switched on at $t=0$,

$$\langle\rho(t, f)\rangle_{Av^{(s)}}=2\int_0^t\langle G(t)G(t-\tau)\rangle_{Av^{(s)}}\cos 2\pi f\tau d\tau$$

$$=2\int_0^t\Psi(\tau)\cos 2\pi f\tau d\tau \tag{19}$$

and the stochastic average* of the instantaneous power spectrum asymptotically approaches the function $P(f)$. The result (19) shows how the power spectrum of the process develops in time. At time t after starting the process, if t is sufficiently large that the process has effectively "forgotten" its start, i.e., $\Psi(\tau>t)\simeq 0$, the instantaneous power spectrum has effectively attained its ultimate value. This means that a sufficient statistical sample of the process has been taken for estimating the power spectrum.

CONCLUSIONS

It appears that an instantaneous power spectrum can be rigorously defined as a useful concept, and that its properties not only fill an intuitive need, but also serve to round out the mathematical structure of spectrum analysis.

[1] See S. O. Rice, Bell System Tech. J., 283–332 (July 1944).

* Interchanging the order of stochastic averaging and integration is justified by Fubini's theorem.

The Spectral Emissivity of Iron-Nickel Alloys

H. B. Wahlin, Robert Zentner, and James Martin
Department of Physics, University of Wisconsin, Madison, Wisconsin
(Received September 26, 1951)

The spectral emissivity for $\lambda = 6700$A has been determined for Fe-Ni alloys and shows changes with temperature which indicate that in the temperature range of 1200°K to 1600°K abnormal behaviors exist which hitherto have not been suspected. A correlation of emissivity change with change in density of metals is suggested.

In two earlier reports[1,2] results have been presented showing that in the case of iron changes in emissivity take place at the A_3 and A_4 points and in cobalt in the Curie temperature range. In Fe-W alloys and Fe-Co alloys changes were found at temperatures where phase changes occur.

The present report is an extension of the measurements to the Fe-Ni system. The technique used in these measurements is essentially the same as that used in previous work. Briefly it consists in measuring the true and apparent temperatures, using a heated hollow cylinder, and computing the emissivity from these two temperatures. The preparation of hollow cylinders has been described earlier.

There is considerable doubt whether pure alloys can be prepared by melting. Since impurities from the crucible used and even some of the material of the crucible itself will dissolve in the molten metal. Some of these impurities, even if present in traces only, might diffuse to the surface of the material during subsequent heating and give spurious values for the emissivity.

In order to avoid this difficulty as far as possible, it was decided to use powder metallurgy techniques in the preparation of the samples. Reagent quality powders were mixed in the desired proportions by passing the mixture repeatedly through a fine sieve. This we believe gives surer mixing of the particles, since it prevents the tendency of the fine powder to cohere in large grains.

The mixed powder was next compressed at room temperature in a hardened tool-steel mold into disks about the size of a half dollar. The pressure used was about 70,000 psi. This pressure was sufficient so that the disks could be removed from the mold without breaking and could be transferred to a furnace for sintering in a flowing hydrogen atmosphere at a temperature of 1100°C. Initially the disks had considerable porosity, but this is an advantage since it allows the hydrogen to remove oxygen and any other impurities which combine with hydrogen from the bulk of the metal.

The initial sintering and heat treatment lasted for 48 hours, although there is no special reason for choosing this time interval. The disks were then quite strong and could be cold-rolled without danger of breaking. The rolling was carried out with frequent periods of annealing at 1000°C in hydrogen until strips $\frac{1}{10}$ mm in thickness resulted. These could then be formed into cylinders for the emissivity measurements.

When the cylinders were first heated in hydrogen, bright pin-point spots were visible on the surface. These were undoubtedly due to miniature cavities in the metal, which gave a blackbody effect. With continued heating these disappeared. This was probably due to the fact that they were filled in by the migration of atoms along the surface.

Some doubt has been raised as to the desirability of preparing these alloys by mixing the powders of the components, since the time needed for interdiffusion might be very long. A calculation of the time needed for complete interdiffusion would be tedious. It was decided therefore to check this experimentally.

In the results on a Fe-Co alloy Knop[2] showed that there exists a region of instability between the temperatures 1213°K and 1292°K. Knop prepared his alloy by electrodeposition, and it seems reasonable to assume that in this case one would have complete alloying. An alloy of Fe-Co and of the same percentage composition as that used by Knop was prepared by the powder mixture method. This was given the same treatment both as regards pressure and heating as the Fe-Ni alloys, and the emissivity was determined. It was found that the temperature limits of the unstable region were the same for this alloy as for the one used by Knop. From this one can conclude with reasonable certainty that at least as far as the surface of the alloy is con-

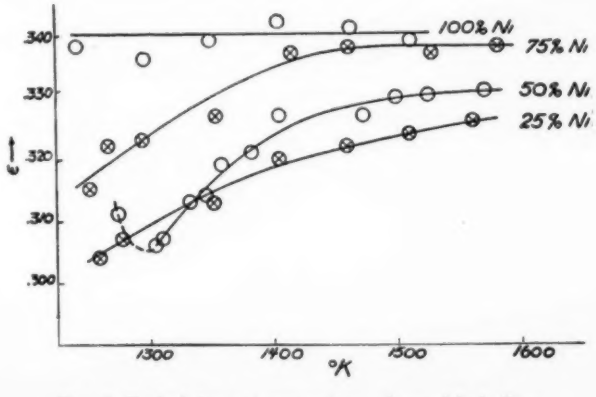


Fig. 1. Emissivity *vs* temperature—iron-nickel alloys.

[1] H. B. Wahlin, and Harry W. Knop, Phys. Rev. **74**, 687–689 (1948).
[2] Harry W. Knop, Phys. Rev. **74**, 1413–1416 (1948).

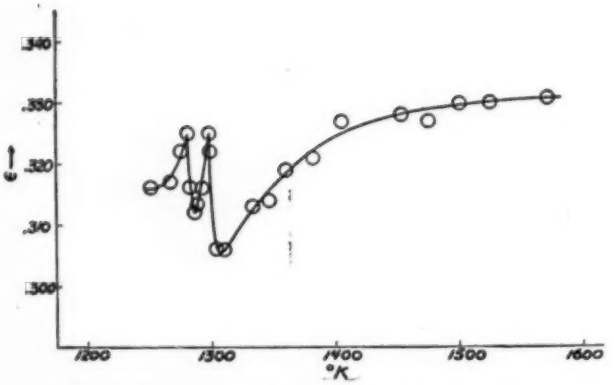


FIG. 2. Emissivity *vs* temperature—50 percent iron, 50 percent nickel.

cerned, complete alloying takes place. One can conclude, in fact, that interdiffusion throughout the entire volume of the alloy is complete, since the specimen prepared from the powders showed the same resistance change with temperature as did Knop's specimen.

RESULTS

Figure 1 shows the emissivity at a wavelength of 6700A as a function of the temperature for a series of composition as indicated. All these show an increase of emissivity with temperature except the 100 percent Ni curve. The 50 percent Ni curve shows an increase at low temperatures and this point could not be brought in line. Figure 2 is a more detailed plot of the 50 percent curve and the increase at low temperatures is shown to be two adjacent peaks. Figure 3 shows the variation of emissivity with temperature for the range 38 percent to 44 percent Ni. This is the invar region where at low temperatures the alloys show abnormally low thermal expansion coefficients. These curves indicate that, contrary to ideas generally held, the abnormal behavior persists at higher temperatures. The alloys from which these curves were obtained were all prepared in this laboratory. An alloy of 41 percent Ni, kindly furnished by Mr. E. M. Wise of the International Nickel Company, fell between the 38 percent and 42 percent curves, and was of the same general shape. In Fig. 3 the values of emissivity as labeled on the ordinate refer to the 38 percent curve. The 42 percent and 44 percent curves fall on top of this curve, and are here displaced an arbitrary distance upward so as to prevent confusion in comparing the curves.

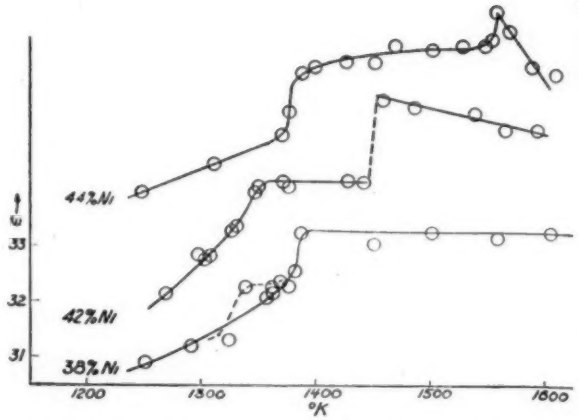


FIG. 3. Emissivity *vs* temperature—iron-nickel alloys.

DISCUSSION

The variation of the emissivity with temperature is still unexplained. In the case of iron there are phase changes at the A_3 and A_4 points. In cobalt there is no phase change in the Curie range, only an expansion of the lattice as shown by Marick.[3] There exists the possibility that the change may be tied up with a change in density of the material, and that this is the principal factor changing the emissivity. Thus when the temperature of iron is raised through the A_3 point, there is an increase in density and a drop in emissivity. Similarily, as the temperature is raised through the A_4 point, there is a drop in the density and an increase in emissivity. The lattice expansion of cobalt observed when the temperature is raised through the Curie range is accompanied by an increase in the emissivity.

In order to determine whether or not these correlations can be extended to the present results, it will be necessary to determine the expansion of these alloys as a function of temperature for the temperature range over which the emissivity was measured. This work is in progress.

In conclusion, the authors wish to express their appreciation of the interest shown by Mr. E. M. Wise of the International Nickel Company. The research was supported by a grant from the Wisconsin Alumni Research Foundation.

[3] Louis Marick, Phys. Rev. **49**, 831–837 (1936).

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 1 JANUARY, 1952

Thermal and Shot Fluctuations in Electrical Conductors and Vacuum Tubes*

SALIM S. SOLOMON
California Institute of Technology, Pasadena, California
(Received August 13, 1951)

A new derivation of Nyquist's equation, relative to the amount of thermal fluctuations generated in an electrical conductor, is presented. A generalization of the new proof to include any arbitrary impedance function shows that the original equation as derived by Nyquist is valid only for physically realizable impedances of the minimum reactance type. This is followed by a new and short derivation of the shot noise formula for temperature limited diodes.

IN 1928 J. B. Johnson[1] showed that "Statistical fluctuation of electric charge exists in all conductors, producing random variation of potential between the ends of the conductor."

Working in conjunction with him, Nyquist[2] was able to show on the basis of the statistical theory of thermodynamics that the thermal noise voltage generated in an impedance Z is given by

$$\langle E^2\rangle_{Av}=4RKT\Delta f, \tag{1}$$

where:

E is the rms value of thermal noise voltage,
R the real part of Z in ohms,
T the absolute temperature in degrees absolute,
K Boltzmann's constant 1.37×10^{-23} watt sec per degree Ab,
Δf band width in cycles per second.

Combining electric circuit theory with the equipartition theorem from thermodynamics a simple derivation of Nyquist's equation is possible.

Consider the simple $r-c$ circuit shown in Fig. 1. If r is the site of an emf whose mean square is $\langle v^2\rangle_{Av}$, then, there must be energy stored in the electric field of magnitude $\frac{1}{2}c\langle V^2\rangle_{Av}$, where $\langle V^2\rangle_{Av}$ is the mean square of the potential drop across c. By the equipartition theorem, in statistical mechanics, the average free energy fluctuation is equal to $KT/2$, but the system has only one degree of freedom, and as there can be no free energy stored in the resistor we may write:

$$\tfrac{1}{2}KT=\tfrac{1}{2}c\langle V^2\rangle_{Av}. \tag{2}$$

From circuit theory it is well known that for a simple $r-c$ circuit the following identity is true:[3]

$$\int_0^\infty Rd\omega=\frac{\pi}{2c}, \tag{3}$$

where R is the real part of the parallel combination of r and c. Substituting from Eq. (2) for c in Eq. (3) we get

$$\langle V^2\rangle_{Av}=4KT\int_0^\infty Rdf. \tag{4}$$

This is Nyquist's equation in its integral form. In Eq. (4), if we let $c\to0$, then $R\to r$ a pure resistance independent of frequency, and we may write

$$\langle V^2\rangle_{Av}=\langle v^2\rangle_{Av}=4KTr\int_0^\infty df. \tag{5}$$

Equation (5) has been derived subject to the assumptions of statistical mechanics as related to random phenomena and therefore suffers from all its limitations. It will be explained in a later part of this paper why there is a limiting frequency, above which Eq. (5) cannot be expected to hold.

In computing the integral in Eq. (5) we have to limit ourselves to that part of the spectrum Δf in which the validity of Eq. (5) may be ascertained. Thus the nonconvergence of the integral in Eq. (5) is excluded. To avoid any ambiguity in the interpretation of Eq. (5) it is well to write it in the form

$$\langle v^2\rangle_{Av}=4KTr\Delta f. \tag{6}$$

From a practical point of view the preceding restriction presents no problem since all known resistances have associated with them a capacitance large enough to make the integrand in Eq. (4) vanish long before the limiting frequency is reached.

Generalized Proof of Nyquist's Equation

Let $\theta=A(\omega)+jB(\omega)$ represent any physically realizable impedance function‡ as shown in Fig. 2, where C

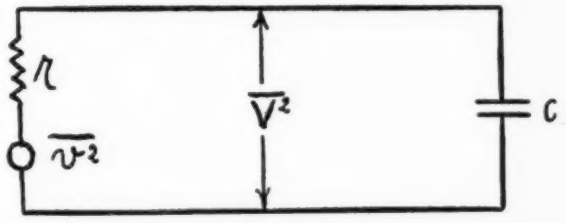


FIG. 1.

* Part of a thesis submitted by the author to the Graduate School at the California Institute of Technology in partial fulfillment of the requirements for the degree of Electrical Engineer.
† Now at the University of California at Berkeley.
[1] J. B. Johnson, Phys. Rev. **32**, 97 (1928).
[2] H. Nyquist, Phys. Rev. **32**, 110 (1928).
[3] H. W. Bode, *Network Analysis and Feedback Amplifier Design* (D. Van Nostrand Company, Inc., New York, 1945).

‡ For a complete treatment of the restrictions placed on θ by the requirements of physical realizability see reference 3.

represents all the shunt capacity that may be pulled out between the terminals $a-b$, A_1+jB_1, any physically realizable impedance, v the value of the voltage generated within A_1+jB_1 and V, the potential drop across $a-b$.

By the Fourier integral energy theorem[4] the energy E between the terminals $a-b$ may be written as

$$E \propto \int_{\infty}^{\infty} [V(t)]^2 dt \propto \frac{1}{\pi}\int_0^{\infty} |V(\omega)|^2 d\omega, \tag{7}$$

where $V(\omega)$ is the Fourier transform of the voltage drop across $a-b$. By simple circuit theory we may reduce Eq. (7) to§

$$E=\frac{\alpha}{\pi}\int_0^{\infty} |v(\omega)|^2 \frac{A}{A_1} d\omega, \tag{8}$$

where α is a proportionality constant and $v(\omega)$ the Fourier transform of the voltage generated within A_1+jB_1.

In order to proceed further with our derivation we have to make some speculation as to the nature of the

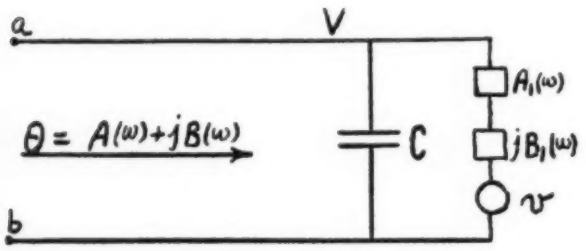


FIG. 2.

voltage that may arise at the terminals of a physically realizable impedance in equilibrium. Fortunately this has been done at the beginning of the twentieth century by many who were trying to explain the phenomena of Brownian motion and random disturbances in general through the application of the theory of statistical mechanics. An excellent treatment is given by Lorentz[5] from whom we choose to quote: "The random motion of electrons, similar to the thermal agitation of gas molecules will give rise to spontaneous electric currents whose direction and intensity vary continuously." By treating the random motion of electrons in a simple $r-L$ circuit as constituting an impulse, Lorentz has virtually derived Nyquist's equation in the time domain long before Nyquist published his results. Thus we feel justified in assuming that $v(t)$ is the resultant of a large number of independent impulses. Furthermore, if we wish to give a theory amenable to experimental test, we must consider the mean square voltage caused by a random disturbance instead of the instantaneous emf. In reality the emf alters its direction so often that the observed voltage is merely a smoothed-over version of the actual voltage.

For a system in equilibrium the mean square value $\langle v^2\rangle_{Av}$ of a disturbance $v(t)$ is defined as

$$\langle v^2\rangle = \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}[v(t)]^2 dt. \tag{9}$$

Equation (9), when written in terms of the Fourier transform of $v(t)$, becomes[6]

$$\langle v^2\rangle_{Av}=\frac{1}{\pi}\int_0^{\infty}|v(\omega)|^2 d\omega. \tag{10}$$

It is well known that the Fourier transform of an impulse of length τ is a constant independent of frequency for frequencies $f \ll 1/\tau$. Thus Eq. (10) reduces to

$$\langle v^2\rangle_{Av}=\frac{v^2}{\pi}\int_0^{\infty} d\omega, \quad f\ll\frac{1}{\tau}. \tag{11}$$

The restriction $f\ll 1/\tau$ is necessary because at frequencies comparable with $1/\tau$ the phenomena we are considering cease to be random and therefore cannot be expected to obey the laws of statistical mechanics. Thus we have to limit ourselves in computing the integral in Eq. (11) to that part of the spectrum Δf for which $f\ll 1/\tau$. This is a limitation equivalent to that imposed on the integral in Eq. (5), and for the same reason it is well to write Eq. (11) in the form

$$\langle v^2\rangle_{Av}=\frac{v^2}{\pi}\Delta\omega = 2v^2\Delta f. \tag{12}$$

It is of interest to note in this place that the above reasoning is not in conflict with quantum theory, which speculates that there is a limiting frequency above which the result does not hold. According to quantum theory the limiting frequency is given by

$$f=\frac{KT}{h}=6.1\times 10^{12}\text{ cps}, \tag{13}$$

where h is Planck's constant and T the absolute value of room temperature taken as 293°K. In what follows we shall concern ourselves with frequencies much lower than the limiting frequency of 6.1×10^{12} cps.

At frequencies higher than the limiting frequency the energies involved decrease rapidly.

Substituting from Eq. (12) for $|v(\omega)|^2$ in Eq. (8) we have

$$E=\frac{\alpha\langle v^2\rangle_{Av}}{2\pi\Delta f}\int_0^{\infty}\frac{A}{A_1}d\omega. \tag{14}$$

[4] S. Goldman, *Frequency Analysis, Modulation and Noise* (McGraw-Hill Book Company, Inc., New York, 1948).

§ Appendix I.

[5] H. A. Lorentz, "Les Théories Statistiques en Thermodynamique," Conférences faites au Collège de France en Novembre 1912, Leipzig et Berlin, B. G. Teubner Libraire éditeur, (1916).

[6] S. O. Rice, Bell System Tech. J. **23**, 282–332 (1944).

Now from statistical mechanics we know that the total energy per degree of freedom associated with a random disturbance equals KT; also from an energy point of view the terminals $a-b$ in Fig. 2 represent one degree of freedom in the system. The free energy $\frac{1}{2}KT$ associated with this degree of freedom cannot in general be considered to be stored in the condenser alone because A_1+jB_1 may itself include other free energy storage media.

Equating Eq. (14) to KT we get

$$\frac{\alpha\langle v^2\rangle_{Av}}{2\pi\Delta f}\int_0^\infty \frac{A}{A_1}d\omega=KT. \tag{15}$$

To evaluate the integral in Eq. (15), let θ be analytic at zero and infinite frequency so that it may be expanded in power series around zero and infinite frequency as follows:[3]

$$\theta_0=A_0+jB_0\omega+A_1'\omega^2+jB_1'\omega^3+\cdots \tag{16}$$

and

$$\theta_\infty=A_\infty+\frac{jB_\infty}{\omega}+\frac{A_1''}{\omega^2}+\frac{jB_1''}{\omega^3}+\cdots. \tag{17}$$

Now the $\int_0^\infty(A/A_1)d\omega$ may be computed by evaluating the line integral of $(\theta-A_\infty)/A_1$ in the right half of the complex frequency plane, along a contour bounded by a large semicircular arc near infinity and the real frequency axis between ∞ and $-\infty$.

If θ is a minimum reactance impedance then it will contain no singularities within the path of integration[3] and we may write

$$\oint\frac{\theta-A_\infty}{A_1}d\omega=0. \tag{18}$$

Carrying through the integration we find that

$$\int_0^\infty\frac{A-A_\infty}{A_1}d\omega=-\frac{\pi B_\infty}{A_{1\infty}} \tag{19}$$

for our circuit $A_\infty=0$, $B_\infty=-(1/c)$, and $A_{1\infty}=r$, the value of the pure resistance that is contained within A_1+jB_1; substituting from Eq. (19) into Eq. (15) we obtain

$$\frac{\alpha}{c}\frac{\langle v^2\rangle_{Av}}{4r\Delta f}=KT. \tag{20}$$

To evaluate α we let $A_1\to r$, a pure resistance, and our circuit reduces to the simple $r-c$ circuit shown in Fig. 1 for which we have already proved that $\langle v^2\rangle_{Av}=4KTr\Delta f$; thus we deduce $\alpha/c=1$ and obtain the expected result

$$\langle v^2\rangle_{Av}=4KTr\Delta f. \tag{21}$$

Equation (21) definitely associates the noise voltage fluctuations with a physical resistance that may exist in any two terminals, physically realizable impedance of the minimum reactance type.

To show that Eq. (21) is not a general result let us evaluate the integral in Eq. (15) when θ represents a physically realizable impedance of the nonminimum reactance type. Following steps similar to those followed in evaluating Eq. (19) we may prove that¶

$$\int_0^\infty\frac{A-A_\infty}{A_1}d\omega<-\frac{\pi}{2}\frac{B_\infty}{A_{1\infty}}; \tag{22}$$

substituting back in Eq. (15) we get

$$\frac{\alpha}{c}\frac{\langle v^2\rangle_{Av}}{4r\Delta f}>KT; \tag{23}$$

remembering that $\alpha/c=1$ we deduce the interesting result that

$$\langle v^2\rangle_{Av}>4KTr\Delta f. \tag{24}$$

Therefore we may generalize Nyquist's equation and write

$$\langle v^2\rangle_{Av}\geqslant 4KTr\Delta f, \tag{25}$$

where the equality sign holds for all physically realizable impedances of the minimum reactance type.

It is worthwhile to point out in this place that the preceding derivation assumes, as a medium of free energy storage, a system of one degree of freedom, whereas Nyquist in his original proof makes use of a lossless transmission line having an infinite number of degrees of freedom; thus Nyquist's calculations are apt to be complicated without adding any generality to the result. The energy storage medium is used only as a means towards an end; the final result, as it should, is found to be independent of the energy storage medium. Another advantage of the preceding method is that it is capable of application to other systems as is demonstrated in the second part of this paper where the shot noise formula for temperature limited diode is developed using a similar line of approach.

II. SHOT NOISE[7]

In an ordinary vacuum tube, the electric current emitted from a hot cathode consists of the combined effect of a large number of independently emitted electrons. In 1918, W. Schottky[8] described the nature of noise which should theoretically be associated with the random emission of the electron convection current. The magnitude of the noise at low frequency has been calculated by various methods. A new and simple derivation of the shot noise formula for the temperature limited diode will now be described.

Let the anode current be formed of identical independent pulses; let $i(t)$ be the current pulse caused by

¶ Similar integrals are evaluated in more detail in reference 3.

[7] S. S. Solomon, Seminar report, filed at the University of California Engineering Library at Berkeley.

[8] W. Schottky, Ann. Physik **57**, 541–67 (1918).

the random flight of the electron between cathode and anode; let $i(\omega)$ be the Fourier transform of $i(t)$. If transit time effects are to be neglected $i(t)$ represents an impulse and[4]

$$i(\omega)=\lim_{t\to 0}\int_{-\infty}^{\infty} i(t)e^{-j\omega t}dt=e \text{ (the electronic charge).} \quad (26)$$

Also the Fourier energy theorem enables us to write

$$\langle i^2\rangle_{Av}=\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}[i(t)]^2dt$$

$$=\frac{1}{2\pi}\int_{-\infty}^{\infty}|i(\omega)|^2d\omega=\frac{e^2}{\pi}\int_0^{\infty}d\omega. \quad (27)$$

Here again there exists a limiting frequency above which Eq. (27) does not hold, and we must take care in applying it only to that part of the spectrum Δf in which it is valid. In other words it is well to write Eq. (27) in the form

$$\langle i^2\rangle_{Av}=\frac{e^2}{\pi}\Delta\omega=2e^2\Delta f. \quad (28)$$

Equation (28) gives the mean square fluctuation $\langle i^2\rangle_{Av}$ caused by the random flight of one electron, and if there are N such electrons per second moving at random between cathode and anode, we may write

$$N\langle i^2\rangle_{Av}=\langle I^2\rangle_{Av}=2Ne\cdot e\Delta f. \quad (29)$$

It is clear that $Ne=I_0$ the dc anode current. Therefore

$$\langle I^2\rangle_{Av}=2I_0e\Delta f. \quad (30)$$

The preceding derivation is important in that it may be readily generalized to give the amount of noise under any operating conditions in which the current pulses are a given function of time. Thus we may write in general

$$\langle I^2\rangle_{Av}=\frac{\alpha}{\pi}\int_0^{\infty}|i(\omega)|^2d\omega, \quad (31)$$

where α is a parameter of the system.

To illustrate the usefulness of Eq. (31) let us consider the temperature limited diode when transit time becomes important. The current pulse may be easily shown to be[9] given by

$$i(t)=(2e/\tau^2)t, \quad 0<t<\tau,$$
$$i(t)=0 \quad \text{for all other } t; \quad (32)$$

[9] G. E. Duvall, Technical report No. 82, Research Laboratory of Electronics, M.I.T. (1948).

e is the charge of the electron; τ the cathode to anode transit time.

Hence we calculate

$$|i(\omega)|^2=\frac{4e^2}{2\pi\theta^4}\{\theta^2+2(1-\cos\theta-\theta\sin\theta)\}, \quad (33)$$

where $\theta=\omega\tau$.

Therefore

$$\langle I^2\rangle_{Av}=\frac{4e^2\alpha}{2\pi^2}\int_0^{\infty}\frac{1}{\theta^4}\{\theta^2+2(1-\cos\theta-\theta\sin\theta\}d\omega. \quad (34)$$

The constant of proportionality is computed by letting θ go to zero and the power approach its low frequency value: $2eI_0\Delta f$. Hence we deduce that in the frequency band Δf

$$\langle I^2\rangle_{Av}=2eI_0\Delta f\times\frac{4}{\theta^4}\{\theta^2+2(1-\cos\theta-\theta\sin\theta)\}. \quad (35)$$

ACKNOWLEDGMENT

The author would like to express his sincere thanks and gratitude to Professor W. H. Pickering for his helpful suggestions and continuous encouragement to him.

Also, special thanks go to Dr. P. G. Thiene for his constructive and friendly criticism of this paper.

APPENDIX I

Referring to Fig. (2) we may write

$$V(\omega)=\frac{v(\omega)}{j\omega c\left(A_1+jB_1+\dfrac{1}{j\omega c}\right)}; \quad (1)$$

$$|V(\omega)|^2=\frac{|v(\omega)|^2}{[1-\omega cB_1]^2+[\omega cA_1]^2}; \quad (2)$$

but

$$\theta=A(\omega)+jB(\omega)=\frac{A_1+jB_1}{j\omega c\left[A_1+jB_1+\dfrac{1}{j\omega c}\right]}; \quad (3)$$

$$A(\omega)=\frac{A_1}{[1-\omega cB_1]^2+[\omega cA_1]^2}; \quad (4)$$

substituting from (4) in Eq. (2) we get

$$|V(\omega)|^2=|v(\omega)|^2\frac{A(\omega)}{A_1(\omega)}. \quad (5)$$

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 1 JANUARY, 1952

Heat Transfer in Isotropic Turbulence

STANLEY CORRSIN
Department of Aeronautics, The Johns Hopkins University, Baltimore, Maryland
(Received August 6, 1951)

The postulated problem of heat transfer in a stationary isotropic turbulence under constant small temperature gradient is examined from both Lagrangian and Eulerian points of view. By combination of Taylor's "diffusion by continuous movements" and a new temperature fluctuation equation somewhat like the well-known turbulent energy equation, it is possible to obtain (1) an expression for the turbulent heat transfer coefficient which depends only on velocity field, and (2) an approximate expression for the corresponding correlation coefficient which depends only on velocity field and fluid physical properties.

A preliminary result is given for "turbulent Prandtl Number" in a homogeneous shear flow in terms of the shear and heat transfer correlation coefficients.

INTRODUCTION

THERE apparently exists no statistical theory of turbulent heat transfer at the present time, perhaps largely because in practice heat transfer occurs in turbulent shear flows, for which there is no theory. However, some idea of the nature of the transfer process can be gained from a highly simplified configuration: steady state homogeneous heat transfer in a non-decaying isotropic turbulence. It is assumed that the mean temperature gradient, $d\bar{T}/dy$, is constant over a region very large compared with any dimensions characterizing the fluctuation field, and that the density is constant. The latter restriction is intended to correspond to a gas with vanishingly small temperature gradient. For an ideal liquid, c_v replaces c_p throughout.

The analysis will first be done with a specific initial condition especially suited to a Lagrangian analysis following G. I. Taylor's theory of "diffusion by continuous movements."[1] This arrangement approaches the desired configuration asymptotically, i.e., for very large time (t) in the absence of a mean velocity ($\bar{U}$), for very large x in the presence of a mean velocity. This asymptotic state is then examined in terms of Eulerian variables, and the two analyses are combined.

The objective of a turbulent heat transport theory at present should probably be to permit prediction of the transport rate from (*given*) complete statistical information on the turbulence field and complete information on the mean thermal boundary conditions.

Of course, the discussion also applies to turbulent mass transfer in cases with constant molecular diffusion coefficient.

LAGRANGIAN TREATMENT WITH ZERO MOLECULAR CONDUCTIVITY

The Mean Temperature Field

We generalize Taylor's approach[1] to the case of a distributed heat source in the plane $x=0$, with absolute temperature $\bar{T}_0$ given as a linear function of y,

$$\bar{T}_0=\bar{T}_1+(d\bar{T}_0/dy)\cdot y. \qquad (1)$$

We postulate a *nondecaying* isotropic turbulence flowing with mean velocity $\bar{U}$ in the x-direction. The turbulence level is chosen low enough to permit an approximate Lagrangian analysis with x as independent variable in place of the rigorous variable t, the time.[2] This is largely for ease of visualization and to simplify later comparison with the Eulerian analysis.

Using the empirical result that the mean thermal wake behind a line source has simple geometrical similarity,[3,4] we can show that $\bar{T}(x, y)$ is independent of x, i.e., equal to $\bar{T}_0(y)$. In a sense this may be considered obvious from conservation of heat and the postulate of homogeneity.

Consider each differential source strip of width dy as a line source whose mean thermal wake has temperature field

$$\overline{\Theta}_1(x, y, y_0)=\overline{\Theta}_m(x, y_0)\phi[(y-y_0)^2/s^2(x)], \qquad (2)$$

where s is the standard deviation, y_0 is the local source location in the plane $x=0$.

From conservation of heat, for given y_0,

$$\overline{\Theta}_m\cdot s=\text{constant}=h\Big/\rho c_p\bar{U}\int_{-\infty}^{\infty}\phi d\eta; \quad \eta=(y-y_0/s),$$

h is total rate of heat generation by the element at $x=0$. With $\overline{\Theta}_m$ a linear function of y_0, we see that

$$\overline{\Theta}_m(x, y_0)=[H/s(x)]\{\bar{T}_1+(d\bar{T}_0/dy)\cdot y_0\}. \qquad (3)$$

H is a constant.

The mean temperature fields of any two sources in such a diffusion problem are obviously superposable, since the heat transfer equation is linear in temperature. Expressed more physically, the mean temperature level at a point is proportional to the rate at which heat pulses cross the point; since heat pulses from different sources have no mutual interference, their rates of passage at a given point are simply additive. Extending

[1] G. I. Taylor, Proc. London Math. Soc. **A20**, 196–211 (1922).

[2] G. I. Taylor, Proc. Roy. Soc. (London) **A151**, 421–478 (1935).

[3] D. C. Collis, Report No. A. 55, Div. of Aeronaut., Australian Council for Sci. and Indust. Research (1948).

[4] M. S. Uberoi and S. Corrsin, Report submitted to NACA for contract NAw-5804 (to be published).

this to the continuous case, we calculate the temperature field of the distributed source by integration over y_0;

$$\bar{T}(x, y)=\int_{-\infty}^{\infty} \overline{\Theta}_1(x, y, y_0)dy_0. \tag{4}$$

With Eqs. (2) and (3), and the variable η, this becomes

$$\bar{T}(x, y)=H\{\bar{T}_1+(d\bar{T}_0/dy)\cdot y\}\int_{-\infty}^{\infty}\phi d\eta - Hs\int_{-\infty}^{\infty}\eta\phi d\eta.$$

Because of the symmetry of ϕ, the last term is zero:

$$\bar{T}(x, y)=H\{\bar{T}_1+(d\bar{T}_0/dy)\cdot y\}\int_{-\infty}^{\infty}\phi d\eta,$$

which is independent of x. For consistency with the boundary condition at $x=0$,

$$H=\left[\int_{-\infty}^{\infty}\phi d\eta\right]^{-1},$$

and

$$\bar{T}(x, y)=\bar{T}_1+(d\bar{T}_0/dy)\cdot y \tag{5}$$

which shows that the mean temperature field is independent of x, for $x>0$. For $x<0$ the temperature is, of course, constant.

The Temperature Fluctuations

Consider a fixed point ($x>0$) in the flow described above. In the absence of molecular conductivity, the instantaneous temperature at the point will be given uniquely by the y coordinate at which the fluid "particle" passed through the plane $x=0$ at some time τ in the past, since there are no temperature fluctuations for $x<0$. With our postulate of small turbulence level, $\tau\approx x/\bar{U}$. Hence we set up a reversed "diffusion by continuous movements" analysis, considering the previous dispersion of particles which later pass through the chosen point, instead of considering the dispersion of all particles which originally passed through a fixed point. With nondecaying turbulence, the dispersion analysis is symmetrical in time, and Taylor's ideas can be used directly.

Let

$$Y(t)=\int_0^t v(t)dt$$

be the y-distance traveled by a fluid particle in time t. Any particle which has traveled such a distance between the plane $x=0$ and passage through our fixed point of observation brings with it a temperature perturbation (deviation from local $\bar{T}$),

$$\vartheta(t)=-Y(t)(d\bar{T}/dy). \tag{6}$$

The negative sign signifies a negative temperature fluctuation when $(d\bar{T}/dy)>0$ and $Y>0$. Equation (6) gives an immediate expression for rms temperature fluctuation at a point,

$$\vartheta'=Y'(d\bar{T}/dy), \tag{7}$$

independent of y. $\vartheta'^2\equiv\langle\vartheta^2\rangle_{Av}$; $Y'^2\equiv\langle Y^2\rangle_{Av}$. $Y'(t)$, or $Y'(x)$, is given by previous work on diffusion from local sources:[1, 5]

$$Y'^2(t)=2v'^2\int_0^t (t-\tau)R_v(\tau)d\tau, \tag{8a}$$

or, approximately,

$$Y'^2(x)=2(v'/\bar{U})^2\int_0^x (x-\xi)R_v(\xi)d\xi, \tag{8b}$$

where

$$v'^2R_v(\tau)=\langle v(t)v(t-\tau)\rangle_{Av},$$

the Lagrangian correlation coefficient for the velocity of a given fluid particle, as introduced by Taylor.

We observe from (7) that without molecular conductivity the temperature fluctuation level increases indefinitely with t or x. From the well-known limiting behaviors of Y', we find that for $x\to 0$,

$$\vartheta'(x)=(v'/\bar{U})(d\bar{T}/dy)\cdot x, \tag{9}$$

while for $x\to\infty$,

$$\vartheta'(x)=(v'/\bar{U})(d\bar{T}/dy)(2L_vx)^{\frac{1}{2}}, \tag{10}$$

where L_v is a Lagrangian scale,

$$L_v=\int_0^\infty R_v(\xi)d\xi.$$

The Turbulent Heat Transport

Our primary objective is investigation of the turbulent heat transport, which is proportional to the (ϑv)-correlation. From (6),

$$\langle\vartheta v\rangle_{Av}=-\langle Yv\rangle_{Av}(d\bar{T}/dy). \tag{11}$$

But

$$\langle Yv\rangle_{Av}=\langle Y(dY/dt)\rangle_{Av}=\tfrac{1}{2}(d/dt)\langle Y^2\rangle_{Av},$$

for which Taylor deduced an expression in terms of Lagrangian correlations coefficient. Thus,

$$\langle\vartheta v\rangle_{Av}=-v'^2(d\bar{T}/dy)\cdot\int_0^t R_v(\tau)d\tau, \tag{12a}$$

or, in the flowing system, we have approximately

$$\langle\vartheta v\rangle_{Av}=-(v'^2/\bar{U})(d\bar{T}/dy)\int_0^x R_v(\xi)d\xi. \tag{12b}$$

For $x\to 0$, $R_v\to 1$ and

$$\langle\vartheta v\rangle_{Av}=-(v'^2/\bar{U})(d\bar{T}/dy)\cdot x. \tag{13}$$

[5] J. M. Kampé de Feriet, Ann. Soc. Sci. Bruxelles, **59**, 145–194 (1939).

For $x\to\infty$,

$$\int_0^x R_v d\xi \to L_v,$$

and

$$\langle\vartheta v\rangle_{kn} = -(v'^2/\bar{U})(d\bar{T}/dy)\cdot L_v \tag{14}$$

which is independent of position.

The heat transfer correlation coefficient,

$$R_{\vartheta v} \equiv \langle\vartheta v\rangle_{kn}/\vartheta' v',$$

decreases monotonically with x:

$$R_{\vartheta v} = -\left\{2\int_0^x (x-\xi)R_v(\xi)d\xi\right\}^{-\frac{1}{2}}\cdot\int_0^x R_v(\xi)d\xi. \tag{15}$$

For $x\to 0$,

$$R_{\vartheta v} = -1. \tag{16}$$

For $x\to\infty$,

$$R_{\vartheta v} = -(L_v/2x)^{\frac{1}{2}} \to 0. \tag{17}$$

For a suitable $x\to\infty$ asymptotic behavior of this problem we might feel that both $R_{\vartheta v}$ and $\langle\vartheta v\rangle_{kn}$ should become nonzero constants. It appears that the continuous increase in $\vartheta'(x)$ has precluded this possibility, and therefore we shall be forced to include the effect of molecular conduction in putting a bound on the value of ϑ'.

It is interesting to observe that $\langle\vartheta v\rangle_{kn}$, effectively the turbulent heat transport, does have a finite asymptote, even with the unbounded increase in ϑ' and constant v'. This is apparently because molecular heat conduction has no direct effect on turbulent heat transfer. It suggests that Eq. (12) may actually represent the turbulent heat transport even in a conducting fluid, as will be discussed later.

Before passing to the Eulerian analysis, it may be pertinent to consider other forms of (14). The turbulent heat transport rate is

$$Q_T = -\rho c_p\langle\vartheta v\rangle_{kn}. \tag{18}$$

For $x\to\infty$, we may introduce a transverse Lagrangian scale, $\mathfrak{L} = (v'/\bar{U})\cdot L_v$ whence

$$Q_T = \rho c_p v'\mathfrak{L}(d\bar{T}/dy). \tag{19}$$

If we choose to define the turbulent "exchange" or "diffusion" coefficient k_T in a conventional way,

$$Q_T = k_T(d\bar{T}/dy),$$

it follows that

$$k_T = \rho c_p v'\mathfrak{L}, \tag{20}$$

and we observe that the characteristic length $\mathfrak{L}$ apparently plays a role much like that assigned to "mixing length" in the older semi-empirical transport analyses. Similar conclusions have been reached in reference 4, where heat transport from a line source in isotropic turbulence was studied. $\mathfrak{L}$ is the same as a length (l_1) introduced for other reasons by G. I. Taylor[2] into the analysis of diffusion from a local source in a decaying turbulence.

This dependence of turbulent transport upon Lagrangian scale is quite reasonable in view of the fact that the scale is simply the zero-wave-number intercept (="dc component") of the corresponding one-dimensional power spectrum.

EULERIAN ANALYSIS

Here we examine the field from an Eulerian point of view at very large t or x. Hence we postulate a nondecaying isotropic turbulence with given mean temperature gradient $d\bar{T}/dy$.

The mean, steady state turbulent heat transfer equation for a constant density fluid (apparently first given by Kampé de Feriet[6]) is

$$\bar{U}_k(\partial\bar{T}/\partial x_k) = \gamma(\partial^2\bar{T}/\partial x_j\partial x_j) - (\partial/\partial x_i)\langle\vartheta u_i\rangle_{kn} \qquad (\gamma\equiv k/\rho c_p) \tag{21}$$

where the repeated index indicates summation. For this particular problem it reduces to the obvious statement that

$$(\partial/\partial y)\langle\vartheta v\rangle_{kn} = 0. \tag{22}$$

However, it is possible to deduce a new equation, one which describes the history of the temperature fluctuations. This follows from the general heat transfer equation in the same way as the well-known turbulent energy equation follows from the Navier-Stokes equations:

With a Reynolds-type hypothesis for the temperature $[T(t) = \bar{T} + \vartheta(t)]$ as well as velocity, the heat transfer equation becomes

$$(\partial\vartheta/\partial t) + (\bar{U}_k + u_k)(\partial/\partial x_k)(\bar{T}+\vartheta) = \gamma(\partial^2/\partial x_i\partial x_i)(\bar{T}+\vartheta). \tag{23}$$

Subtracting (21) from (23), we get the equation for the temperature fluctuations,

$$\frac{\partial\vartheta}{\partial t} + \bar{U}_i\frac{\partial\vartheta}{\partial x_i} + u_j\frac{\partial\bar{T}}{\partial x_j} + \frac{\partial}{\partial x_k}(\vartheta u_k) - \frac{\partial}{\partial x_l}\langle\vartheta u_l\rangle_{kn} = \gamma\frac{\partial^2\vartheta}{\partial x_m\partial x_m}. \tag{24}$$

This is multiplied by ϑ and averaged, giving the final equation for $\langle\vartheta^2\rangle_{kn}$:

$$\frac{\bar{D}}{Dt}\langle\vartheta^2\rangle_{kn} = -2\langle\vartheta u_k\rangle_{kn}\frac{\partial\bar{T}}{\partial x_k} - \frac{\partial}{\partial x_m}\langle u_m\vartheta^2\rangle_{kn} + \gamma\frac{\partial^2}{\partial x_j\partial x_j}\langle\vartheta^2\rangle_{kn} - 2\gamma\left\langle\frac{\partial\vartheta}{\partial x_i}\frac{\partial\vartheta}{\partial x_i}\right\rangle_{kn}, \tag{25}$$

where $\bar{D}/Dt$ denotes mean "Stokes derivative"

$$[(\partial/\partial t) + \bar{U}_i(\partial/\partial x_i)].$$

[6] J. M. Kampé de Feriet, Ann. Soc. Sci. Bruxelles, **57**, 67–72 (1937).

Equation (25), valid for incompressible turbulent heat transfer in general, clearly is much the same form as the turbulent energy equation. Of course the lack of perfect analogy shows up here in the absence of a term corresponding to the pressure-velocity correlation term of the turbulent energy equation. Equation (25) says that the time rate of change of temperature fluctuation "power" in a volume which follows the *mean* fluid motion is equal to (a) the rate at which new fluctuations are produced by the action of the turbulence on the mean temperature gradient plus (b) the rate at which temperature fluctuation "power" is convected into the volume by turbulent velocity fluctuations plus (c) the last two terms which together describe the molecular conduction of temperature fluctuation "power" into the volume and the rate of destruction of temperature fluctuations by conductive "smearing" within the volume.

For the homogeneous problem under consideration, Eq. (25) reduces to

$$(d\bar{T}/dy)\langle\vartheta v\rangle_{Av}=-\gamma\langle(\partial\vartheta/\partial x_i)(\partial\vartheta/\partial x_i)\rangle_{Av} \qquad (26)$$

which expresses the homogeneous balance between production and "dissipation" of temperature fluctuations.

In order to get a convenient approximation to (26), we postulate the existence of "local isotropy" in the temperature fluctuation field and restrict the discussion to Péclet numbers sufficiently large that virtually all of the "dissipation" takes place in the isotropic range of wave numbers. Kolmogoroff's hypothesis of local isotropy for velocity fluctuations in turbulent shear flow[7] has been well verified by experiment.[8–10]

Under these conditions the right side of Eq. (26) can be approximated by its form for an isotropic field,[11]

$$(d\bar{T}/dy)\langle\vartheta v\rangle_{Av}=-6\gamma(\langle\vartheta^2\rangle_{Av}/\lambda_\vartheta^2), \qquad (27)$$

where λ_ϑ is the Eulerian "microscale" of the temperature fluctuation field. Even though λ_ϑ can be estimated in terms of the turbulence microscale λ (as will be done later) Eq. (27) still leaves us short of the goal of expressing the turbulent heat transfer entirely in terms of mean temperature gradient and the statistics of the velocity field. However, it appears that some reasonable results can be achieved by exploitation of the (conductionless) Lagrangian analysis.

COMBINED ANALYSIS

First let us derive Eq. (14) by a modified method which employs both Lagrangian and Eulerian concepts: for $k=0$, there is a steady increase in ϑ' with x, given by the pertinent special form of Eq. (25):

$$\bar{U}(d\vartheta'^2/dx)=-2\langle\vartheta v\rangle_{Av}(d\bar{T}/dy). \qquad (28)$$

But from the ($k=0$) Lagrangian analysis for $x\to\infty$, Eq. (10) gives

$$(d\vartheta'^2/dx)=2(v'/\bar{U})\mathfrak{L}(d\bar{T}/dy)^2. \qquad (29)$$

Substituting (29) into (28),

$$\langle\vartheta v\rangle_{Av}=-v'\mathfrak{L}(d\bar{T}/dy), \qquad (30)$$

which is identical with (14).

The next step is consideration of the applicability of Eq. (30) to turbulent heat transfer with $k\neq 0$. For this we recall Taylor's discussion[2] on Schubauer's measurements of turbulent diffusion of heat close behind a line source.[12] Taylor pointed out that since the random molecular and turbulent motions (both transporting heat) are statistically independent, the standard deviation of the actual thermal wake was simply the square root of the sum of the squares of the standard deviation which would have resulted from molecular motion alone and of that which would have resulted from turbulent motion alone:

$$s^2=s_{\text{lam}}{}^2+s_{\text{turb}}{}^2.$$

But the diffusion coefficient at any distance from a local source in a random homogeneous field is[4]

$$k\sim ds^2/dx \quad \text{(independent of } y\text{)}.$$

Hence the molecular and turbulent diffusion coefficients add linearly. Of course this simple statement alone can be made directly from Eq. (21) plus the definition of k_T. The significant addition is that because of the statistical independence of molecular and turbulent motions, k_T is independent of k and (30) or (20) applies in general for transport in a nondecaying turbulence.

The ratio of turbulent to molecular heat diffusion coefficient (using $k\approx\rho c_p a l$, from elementary kinetic theory of gases) is

$$k_T/k\approx(v'/a)\cdot(\mathfrak{L}/l), \qquad (31)$$

where a is the root mean square molecular velocity and l is the mean free path. As a numerical example, consider $\mathfrak{L}=0.86$ cm, $v'=15.5$ cm/sec, $\lambda=0.42$ cm [actual values from a case in reference 4]. With $a=5\times10^4$ cm/sec, $l=6\times10^{-6}$ cm, $k_T/k=44$.

It is now possible to deduce the proper (nonzero) asymptotic value of the heat transfer correlation coefficient, $R_{\vartheta v}$, to replace Eq. (17). We simply use Eq. (30) for $\langle\vartheta v\rangle_{Av}$ and obtain the proper equilibrium value of ϑ' from (27) (written explicitly later in this section). From these,

$$R_{\vartheta v}=(6\gamma\mathfrak{L}/\lambda_\vartheta{}^2 v')^{\frac{1}{2}}. \qquad (32)$$

To put this entirely in terms of velocity field (and physical properties of the fluid), the λ_ϑ must be replaced. For isotropic temperature fluctuations in an isotropic turbulence, in two limiting cases of both very high and very low Reynolds and Péclet numbers it has been found that[11]

$$\lambda_\vartheta{}^2=(2/\sigma)\lambda^2, \qquad (33)$$

[7] A. N. Kolmogoroff, Compt. rend. acad. sci. URSS (Doklady) **30**, 301–305 (1941); **32**, 16–18 (1941).
[8] A. A. Townsend, Australian J. Sci. **1**, 161–174 (1948).
[9] S. Corrsin, J. Aeronaut. Sci. **16**, 757–758 (1949).
[10] J. Laufer, J. Aeronaut. Sci. **17**, 277–287 (1950).
[11] S. Corrsin, J. Aeronaut. Sci. **18**, 417–423 (1951).
[12] G. B. Schubauer, NACA Tech. Report, No. 524 (1935).

where λ is the Eulerian microscale introduced by Taylor in 1935,[2] and σ is the Prandtl number, $\sigma=(\nu/\gamma)$, ν=kinematic viscosity. With local isotropy of the temperature fluctuations Eq. (33) may be taken as an estimate. Then

$$R_{\vartheta v}=(3\nu\mathfrak{L}/\lambda^2 v')^{\frac{1}{2}}, \tag{34}$$

or, in terms of the conventional turbulence Reynolds number, $R_\lambda\equiv v'\lambda/\nu$,

$$R_{\vartheta v}=[(3/R_\lambda)\cdot(\mathfrak{L}/\lambda)]^{\frac{1}{2}}. \tag{34a}$$

By definition, $0\leq R_{\vartheta v}\leq 1$, and we can make a numerical estimate again from the example used earlier ($R_\lambda=43.5$): $R_{\vartheta v}=0.38$, a reasonable order of magnitude. The only direct measurements of $R_{\vartheta v}$ to date, in a warm turbulent jet[13] (vastly different conditions, of course) showed maximum values on the order of 0.5 to 0.6. On the other hand, behind a local source in a decaying isotropic turbulence it was computed from mean temperature data to have a maximum closer to unity.[4] However, the indirectness of determination plus the scatter of the original data render these values less certain.

Further speculations can be made from Eq. (34a). First, from the well-known relation between Eulerian scale and microscale,

$$\lambda/L\sim 1/R_\lambda, \tag{35}$$

(deducible either from dimensional reasoning or from Eq. (75) of the paper by von Kármán and Howarth[14]), it follows that

$$R_{\vartheta v}\sim(\mathfrak{L}/L)^{\frac{1}{2}}. \tag{36}$$

From some measurements[4] the (right side) constant of proportionality in (35) can be estimated as about 50. Then (36) becomes

$$R_{\vartheta v}\approx\tfrac{1}{4}(\mathfrak{L}/L)^{\frac{1}{2}}. \tag{37}$$

Unfortunately, no one has yet discovered a relation between Lagrangian and Eulerian scales. Figure 34 of reference 4 indicates a general downward trend of $(\mathfrak{L}/L)$ with increasing R_L, and hence with increasing R_λ since $R_\lambda{}^2\sim R_L$, from (35).

An explicit estimate of the rms temperature fluctuation is of some interest. From Eqs. (27) and (30),

$$\vartheta'=\lambda_\vartheta(d\bar{T}/dy)(v'\mathfrak{L}/6\gamma)^{\frac{1}{2}}, \tag{38}$$

$(v'\mathfrak{L}/\gamma)$ is apparently a significant turbulent Péclet number, $P_\mathfrak{L}$ (say), and we can write

$$\vartheta'=\lambda_\vartheta(d\bar{T}/dy)(P_\mathfrak{L}/6)^{\frac{1}{2}}. \tag{38a}$$

With (33) as estimate for $\lambda_\vartheta(\lambda)$,

$$\vartheta'\approx\lambda(d\bar{T}/dy)(v'\mathfrak{L}/6\nu)^{\frac{1}{2}}. \tag{39}$$

The recent development of the theory of axisymmetric tensors[15,16] suggests a generalized approach to the homogeneous heat transfer problem in isotropic box turbulence ($\bar{U}=0$), since the temperature fluctuation field must have statistical axial symmetry about the direction of the given mean temperature gradient. This approach has led to no results yet; the scalar correlation equation obtainable by combining the heat transfer equations at two different points has four unknowns.

A PRELIMINARY RESULT FOR SHEAR FLOW

Although the concept of diffusion by continuous movements has not yet been applied to a homogeneous turbulent shear flow, one new result relating heat and momentum transport can be deduced with the aid of (25), on a purely Eulerian basis.

Consider an incompressible homogeneous turbulent shear flow with $d\bar{T}/dy$=constant and temperature differences small enough to have a negligible dynamic effect. The total rates of momentum and heat transfer are both constant:

$$\tau=\mu(d\bar{U}/dy)-\rho\langle uv\rangle_{Av}=\text{const}, \tag{40}$$

$$Q=k(d\bar{T}/dy)-\rho c_p\langle\vartheta v\rangle_{Av}=\text{const}. \tag{41}$$

The turbulent energy equation and the corresponding temperature fluctuation Eq. (25) have forms,

$$\langle uv\rangle_{Av}(d\bar{U}/dy)=-\nu\langle\partial u_i/\partial x_k\cdot\partial u_i/\partial x_k\rangle_{Av}, \tag{42}$$

$$\langle\vartheta v\rangle_{Av}(d\bar{T}/dy)=-\gamma\langle(\partial\vartheta/\partial x_i)\cdot(\partial\vartheta/\partial x_i)\rangle_{Av}. \tag{43}$$

With sufficiently high Reynolds and Péclet numbers, local isotropy permits approximating these by

$$\langle uv\rangle_{Av}(d\bar{U}/dy)=-5\nu(q'^2/\lambda^2), \tag{44}$$

$$\langle\vartheta v\rangle_{Av}(d\bar{T}/dy)=-6\gamma(\vartheta'^2/\lambda_\vartheta{}^2), \tag{45}$$

where $q'^2=u'^2+v'^2+w'^2$, the mean square velocity fluctuation.

The transport coefficients for momentum and heat are defined by

$$\epsilon_u=-\langle uv\rangle_{Av}(d\bar{U}/dy)^{-1}, \tag{46}$$

$$\epsilon_\vartheta=-\langle\vartheta v\rangle_{Av}(d\bar{T}/dy)^{-1}, \tag{47}$$

and we would like an expression for the "turbulent Prandtl number,"

$$\sigma_T=\epsilon_u/\epsilon_\vartheta. \tag{48}$$

After substituting for the mean gradients with (44) and (45), we arrive at

$$\sigma_T=(6/5\sigma)\{\vartheta'\langle uv\rangle_{Av}\lambda/q'\langle\vartheta v\rangle_{Av}\lambda_\vartheta\}^2. \tag{49}$$

For a numerical check on the order of magnitude, we insert experimental data for the maximum transport region in a round heated turbulent air jet[13] with assump-

[13] S. Corrsin and M. S. Uberoi, NACA Tech. Note No. 1865, (1949) (reissued as NACA Tech. Report No. 998, (1950)).

[14] Th. v. Kármán and L. Howarth, Proc. Roy. Soc. (London) **A164**, 192–215 (1938).

[15] G. K. Batchelor, Proc. Roy. Soc. (London) **A186**, 480–502 (1946).

[16] S. Chandrasekhar, Phil. Trans. Roy. Soc. (London) **A242**, 557–577 (1950).

tions (a) $w'=v'$ and (b) λ and λ_ϑ are the same as on the jet axis, the only station at which they were measured. Then (49) gives $\sigma_T=0.32$, while the measured value is $\sigma_T=0.6$. The agreement is not good, but if we recall that *a priori* we know only that $0\leq\sigma_T\leq\infty$, the order of magnitude is reasonable. This experimental case is, of course, quite nonhomogeneous.

It is now fairly well established experimentally that σ_T is independent of σ,[17] so it must be possible to eliminate σ from Eq. (49). We recall that in three widely different special cases of isotropic turbulence it has been shown that $\lambda^2\sim\sigma\lambda_\vartheta^2$.[9,18] If we assume that this applies to shear flow with local isotropy, (56) becomes independent of σ:

$$\sigma_T\sim[\vartheta'\langle uv\rangle_M/q'\langle\vartheta v\rangle_M]^2=[(u'/q')\cdot(R_{uv}/R_{\vartheta v})]^2, \quad (50)$$

where $u'v'R_{uv}=\langle uv\rangle_M$.

If we use the particular relation given by (33),

$$\sigma_T=\tfrac{3}{5}[(u'/q')\cdot(R_{uv}/R_{\vartheta v})]^2, \quad (51)$$

and if $u'=v'=w'$,

$$\sigma_T=\tfrac{1}{5}(R_{uv}/R_{\vartheta v})^2. \quad (52)$$

[17] W. Forstall, Jr., and A. H. Shapiro, J. Appl. Mech. **17**, 399–408 (1950).

[18] S. Corrsin, J. Appl. Phys. **22**, 469–473 (1951).

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 1 JANUARY, 1952

X-Ray Measurement of Order in CuPt*

C. B. WALKER
Massachusetts Institute of Technology, Cambridge, Massachusetts

(Received September 20, 1951)

In investigating the order-disorder transformation in CuPt, cell parameters were determined from specimens quenched at intervals from 25°C to 890°C. The rhombohedral deformation of the ordered, layered structure was found to diminish with increasing temperature, disappearing at a critical temperature of 815°C. Integrated intensity measurements of superstructure reflections from quenched powder and briquette samples determined that the long range order, S, decreases slowly with increasing temperature to a value of 0.77 at T_c, where it changes abruptly to zero, with the transformation below T_c being homogeneous. A simple theory is developed relating the rhombohedral deformation to S. Cold working an ordered sample was found to transform the alloy almost completely from the ordered rhombohedral cell to the disordered cubic cell. From measurements of diffuse intensities diffracted by briquettes above T_c, Fourier transform methods determined parameters which fixed the local order as being local layering configurations, in contrast to the usual A-likes-B type of order.

INTRODUCTION

INVESTIGATIONS by Johansson and Linde[1] and Linde[2] of the alloys of Cu with Pt demonstrated ordering transitions around the compositions Cu_3Pt and CuPt. Face-centered cubic in its disordered form, ordering takes place around the CuPt composition below a critical temperature of 815°C by segregation of the Cu and Pt atoms onto alternate (111) planes. This layering deforms the lattice from cubic to rhombohedral symmetry, but since the deformation is only of the order of 1 percent, each atom effectively still has 12 nearest neighbors. Since each atom has 6 like and 6 unlike nearest neighbors in both the ordered and disordered phases, there is no increase in the number of unlike neighbors in the ordering of CuPt. This is quite different from most of the order-disorder transformations, and a further knowledge of the course of this transition becomes of interest.

* Research sponsored by the ONR under Contract N5-ori-07832.

[1] C. H. Johansson, and J. O. Linde, Ann. Physik **82**, 449 (1927).

[2] J. O. Linde, Ann. Physik **30**, 151 (1937).

STRUCTURE DETERMINATION

Samples were taken from three different slugs, all made from C. P. copper shot and platinum wire, rated at less than 0.1 percent impurity, and melted in graphite crucibles in an induction furnace. After homogenizing anneals (1250°C 12 hr for slug No. 1; 1400°C 2 hr for slugs No. 2 and No. 3) Debye-Scherrer photographs showed sharp, clearly resolved $\alpha_1-\alpha_2$ doublets for the high angle lines. Chemical analysis gave for slug No. 1, 49 percent Cu; for slug No. 2, 51 percent Cu; and for slug No. 3, 49 percent Cu, with an estimated accuracy of 1 percent.

To follow the transformation most easily, Johansson and Linde used a 32-atom ordered unit cell, whose axes are the slightly deformed axes of the disordered, face-centered cubic unit cell, doubled in length. The rhombohedral angle of this cell can be described in radian measure as

$$\alpha=\pi/2+\delta,$$

where δ is a small quantity. To determine the cell parameters over the temperature range from 25°C to

890°C, 200 mesh filings from slug No. 1 were sealed in evacuated Vycor tubing, annealed until approximate thermal equilibrium had been established, quenched in water, and used as specimens for Debye-Scherrer photographs in a 19-cm diameter cylindrical camera. These annealing times ranged from 5 hr or less for temperatures above 600°C to periods of over 100 hours for low temperatures. To insure establishment of equilibrium, above 600°C samples were reannealed with annealing times quadrupled, and below 600°C, since this process would have required exorbitant periods, reliance was put on the general behavior of the experimental points. Values of the parameters, δ and a_r, were obtained from the diffraction patterns by the standard methods of analysis; with an estimated accuracy of two parts in the third decimal place of a_r† and an accuracy of 5 percent for δ. Table I compares the cell parameters obtained at two temperatures with those from Linde's data, the latter being extrapolated to a composition of 51 percent Pt and transformed from Kx to A units. The discrepancies between the two are thought to be caused by errors in the analyses of the composition of the alloys.

A plot of these parameters is found in Fig. 1. The

TABLE I. Comparison of cell parameters with those of Linde for the compositon 49 Cu, 51 Pt.

	720°C		300°C	
	a_r(A)	δ(rad)	a_r(A)	δ(rad)
Linde	7.593	0.0130	7.594	0.0170
Exper	7.577	0.0131	7.577	0.0174

rhombohedral deformation decreases slowly with increasing temperature, the cell symmetry changing abruptly to cubic at about 815°C. The scatter of values for δ at the higher temperatures is attributed to difficulties in quenching. In a narrow temperature interval between 815°C and 830°C lines corresponding to both ordered and disordered phases were obtained. This can be interpreted as either a real, two-phase equilibrium region or as the result of incomplete quenching. Differentiation between these possibilities requires a high temperature camera, not available in this investigation. However, throughout the course of this transformation below 815°C only lines pertaining to the ordered phase were found, demonstrating that the major course of this transformation must be classified as homogeneous.

X-RAY DETERMINATION OF LONG-RANGE ORDER

The power per unit length at an x-ray diffraction line from a powder sample is given by

$$P' \sim F^2 m(\text{L.P.}) e^{-2M} A(\theta),$$

where F = the structure factor for the reflection; m = the multiplicity of the reflecting planes; (L.P.) = the

† Wavelengths used: CuKα_1 = 1.5405A; CuKα_2 = 1.5443A.

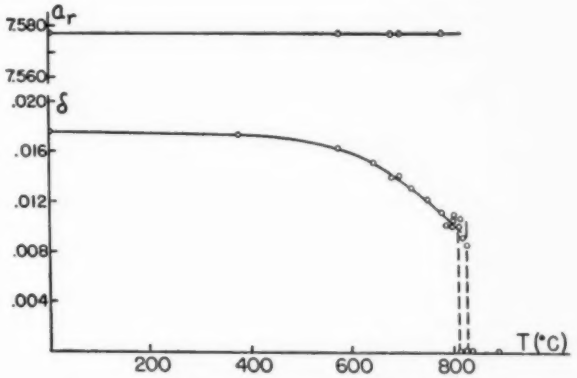


FIG. 1. Cell parameters a_r in A and δ in radians for CuPt quenched from various temperatures.

Lorentz polarization factor; e^{-2M} = the Debye temperature factor; $A(\theta)$ = an absorption correction factor. The structure factor for superstructure reflections is linearly related to the Bragg-Williams order parameter, S, while that for fundamental reflections is independent of S. Thus, if the atomic scattering factors, temperature factors, and absorption correction factors are known, one can determine S as a function of T from experimental measurements of the ratio: P_{sup}'/P_{fund}'. Where these factors are not known accurately, this method cannot be used. If instead measurements are made on quenched specimens and one assumes perfect order at room temperature, then S can be determined simply from measurements of P_{sup}'. For:

$$P_{sup}', T/P_{sup}', 25°\text{C} = S^2. \quad (1)$$

Experimentally the latter method was required, because of inaccurate knowledge of the Hönl L-electron corrections, preferred orientation in pressed briquette samples, and strong, θ-dependent rough surface absorption effects for samples of loose filings. Using a North American Philips Geiger counter spectrometer with filtered CuKα radiation, integrated intensity measurements were made of 5 fundamental and 4 superstructure low angle lines. Samples used were a briquette of filings

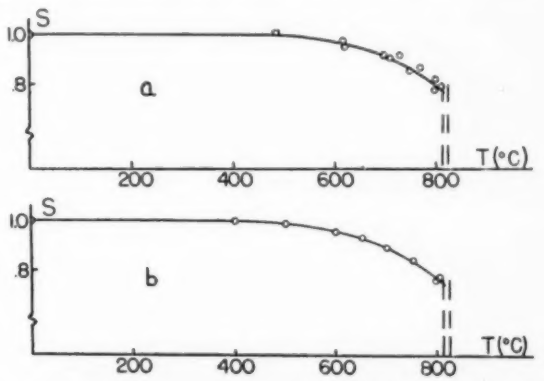


FIG. 2. Long range order S for samples of CuPt quenched from various temperatures. (a) Determined from the intensities of superstructure lines. (b) Calculated from the measured values of δ by Eq. (2).

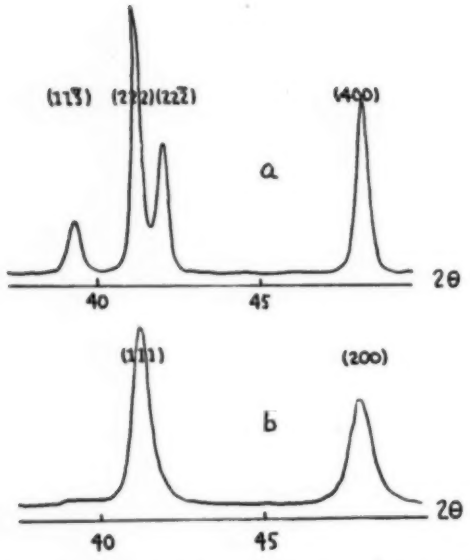


FIG. 3. Powder patterns showing the effect of cold work on an ordered sample of CuPt. (a) Well-ordered and annealed. (b) After cold work. Radiation CuKα.

from slug No. 2, loose 325 mesh filings from slug No. 2, and loose 325 mesh filings from slug No. 3, all annealed in evacuated Vycor tubing and quenched in water, the anneales being longer in each case than those used in determining the cell parameters. To obtain an average surface effect on the loose filing samples, five complete runs were made with the sample from slug No. 2 and three complete runs on the sample from slug No. 3, the peak areas of each run being normalized in terms of its (400) line before averaging.

From these average peak areas values of S^2 *vs* T were obtained from each superstructure line, using Eq. (1). The values of S from the four superstructure lines were averaged, and the resultant values of S *vs* T are shown in Fig. 2a. Long range order is seen to decrease slowly with increasing temperature to a value of about 0.77 at 815°C, where it drops abruptly to zero. Whether this drop is discontinuous or merely very sharp was not determined because of the limitations of quenching techniques. Examination of the integrated intensities of the fundamental lines from these samples showed only small fluctuation, indicating that, within the limitations of quenching, the dependence of S on T was accurate to 5 percent.

LONG-RANGE ORDER AND THE CELL PARAMETERS

Attempts to explain the change in rhombohedral deformation of the ordered cell as a function of quenching temperature by some mechanism other than a varying long range order met with no success, so the development of a relation between S and δ was attempted.

Let: l_1=average distance between two neighboring sites in the same (111) plane; l_2=average distance between two neighboring sites in adjacent (111) planes. By geometry, disregarding quantities of the magnitude of δ^2,

$$(l_1/l_2)^2-1=\tfrac{2}{3}\sqrt{3}\delta.$$

In terms of average nearest neighbor distances, l_{AA}, l_{BB} and l_{AB}, we can write

$$l_1=\tfrac{1}{2}\left\{\begin{matrix} r_B(r_Bl_{BB}+w_Bl_{AB})+w_B(r_Bl_{AB}+w_Bl_{AA}) \\ +r_A(r_Al_{AA}+w_Al_{AB})+w_A(r_Al_{AB}+w_Al_{BB}) \end{matrix}\right\},$$

where r_A, w_A, r_B, w_B are the fractions of A and B sites rightly and wrongly occupied.

Since $S=r_A-w_B=r_B-w_A$,

$$l_1=\tfrac{1}{4}\{(l_{AA}+l_{BB}+2l_{AB})+(l_{AA}+l_{BB}-2l_{AB})S^2\},$$

and similarly,

$$l_2=\tfrac{1}{4}\{(l_{AA}+l_{BB}+2l_{AB})-(l_{AA}+l_{BB}-2l_{AB})S^2\}.$$

Then

$$\left(\frac{l_1}{l_2}\right)^2-1=\frac{S^2}{\{M-(S^2/M)\}^2},$$

where

$$M=\tfrac{1}{2}\left\{\frac{l_{AA}+l_{BB}+2l_{AB}}{l_{AA}+l_{BB}-2l_{AB}}\right\}^{\frac{1}{2}}\gg 1.$$

Neglecting terms in M^{-2},

$$S^2=\frac{\frac{2}{3}\sqrt{3}M^2\delta}{1+\frac{2}{3}\sqrt{3}\delta}. \qquad (2)$$

Assuming M to be independent of the degree of order, it is evaluated by assuming $S=1.0$ at $T=25$°C, where experimentally $\delta=0.0175$ radian. By substituting experimental values of δ in Eq. 2, the plot of S *vs* T found in Fig. 2b is obtained. The agreement with that determined from x-ray intensity measurements is reasonably good.

DESTRUCTION OF ORDER BY COLD WORK

Several early workers have reported that diffraction patterns from cold-worked ordered samples showed marked reduction in superstructure line intensities, the general mechanism being supposedly a random distribution of slippages. Looking for this effect in CuPt, a low angle run was made on a partially ordered briquette with the Philips spectrometer. The briquette was then rubbed by hand on a piece of emery paper, and a second run made over the same angular region. The results are shown by Fig. 3. The (222)−(22$\bar{2}$) fundamental doublet has been essentially transformed into the disordered (111) line, and this line is not broad enough to bear interpretation in terms of a coalescing because of cold-work line broadening. At the same time the (11$\bar{3}$) superstructure line has almost completely disappeared. Thus whatever mechanism is used to explain this phenomenon must account for an almost complete structural change from rhombohedral to cubic symmetry.

SHORT RANGE ORDER IN CuPt

If a binary alloy contains local order configurations but no longer range order, the diffuse intensity in electron units of x-rays diffracted from a powder sample is

$$I_{eu}=Nm_Am_B(f_A-f_B)^2\{1+\sum_i\sum_j\alpha_{ij}(\sin Kr_i/Kr_i)\}, \quad (3)$$

where N = number of atoms in the sample; m_A, m_B, f_A, f_B = atomic fractions and atomic scattering factors of components A and B; $\alpha_{ij}=1-(p_{ij}^A/m_A)=1-(p_{ij}^B/m_B)$; where p_{ij}^A is the probability of finding an A atom in the jth position of the ith shell around any B atom; r_i = radius of the ith shell; and $K=(4\pi\sin\theta/\Lambda)$. The effective coefficient of each modulating $\sin Kr_i$ terms is seen to be $\sum_j\alpha_{ij}=c_i\bar{\alpha}_i$, where c_i is the coordination number of the ith shell. Thus only a spherically averaged $\bar{\alpha}_i$ can be determined experimentally from powder patterns.

Following Wilchinsky[3] let

$$\frac{I_{eu}}{Nm_Am_B(f_A-f_B)^2}-1=\varphi(K),$$

$$c_i\bar{\alpha}_i=\int_{r_i-\epsilon_i}^{r_i+\epsilon_i}f(r)dr,$$

where ϵ_i is chosen so that $f(r)=0$ for the ith shell beyond these limits. Then

$$K\varphi(K)=\int_0^\infty f(r)\frac{\sin Kr}{r}dr,$$

and by a Fourier transform

$$f(r)=\frac{2r}{\pi}\int_0^\infty K\varphi(K)\sin rKdK. \quad (4)$$

Thus if the diffuse intensities can be determined in electron units per atom, a Fourier transform and graphical integration will give the desired parameters, $\bar{\alpha}_i$.

The diffuse intensities scattered by CuPt were first observed in transmission photographs of loose quenched filings, the pictures showing rather sharp diffuse peaks at positions corresponding to the (111) and (11$\bar{3}$) superstructure lines. Further measurements were made on briquette samples in a spectrometer utilizing a G. M. tube detector. With CuKα radiation monochromated by the (10$\bar{1}$1) planes of a bent quartz crystal and focused on the face of the sample, sufficient intensity was available to allow quantitative measurements. The incident and diffracted beams made equal angles with the sample surface, and effects because of $\Lambda/2$ diffraction were removed by use of a Ni–Al balanced filter. Cold-working the faces of the new briquettes with emery paper was found to remove to a large extent the initial preferred orientation, so this treatment was accorded all samples. The contribution of air scattering, fluorescence radiation, Compton scattering, and temperature diffuse scattering to measurements on quenched disordered briquettes was determined by further measurements on the same briquette after annealing below T_c. Subtraction of the two sets of data left the desired diffuse scattering because of local order configurations. This was normalized to electron units per atom by comparison with the high angle scattering of paraffin.

Measurements covering the range from 5° to 25° in θ were made for briquettes quenched from 910°C, 930°C, and 970°C, and for an anneal and quench from 770°C. The resultant diffuse scattering in electron units per atom for CuPt quenched from 930°C is shown in Fig. 4a, the modulations being rather sharp but oscillating about the Laue Monotonic‡ scattering as they should.

To eliminate the question as to whether quenching had retained the high temperature equilibrium distribution, a furnace was designed to measure diffuse intensities at temperatures up to 1000°C. Heating by radiation, one flat wound heater was placed 2 mm behind the sample while an auxiliary heater was placed 20 mm in front to minimize thermal gradients. Oxidation was prevented by the use of a pre-purified nitrogen atmosphere, and temperatures were measured with a chromel-alumel thermocouple bound to the sample

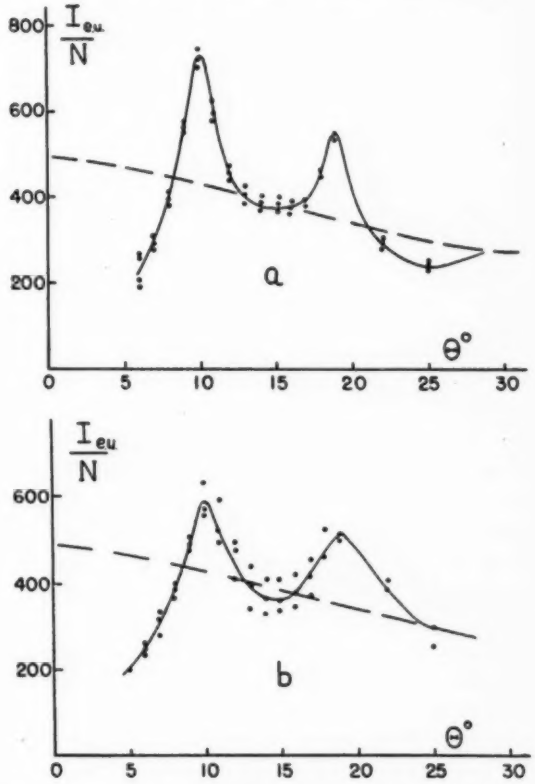


FIG. 4. Short range order diffuse intensity in electron units per atom for CuPt. (a) Sample quenched from 930°C. (b) Measurements on sample held at 890°C.

[3] Z. W. Wilchinsky, J. Appl. Phys. **15**, 806 (1944).

‡ $I_{LM}=Nm_Am_B(f_A-f_B)^2$, see Eq. (3).

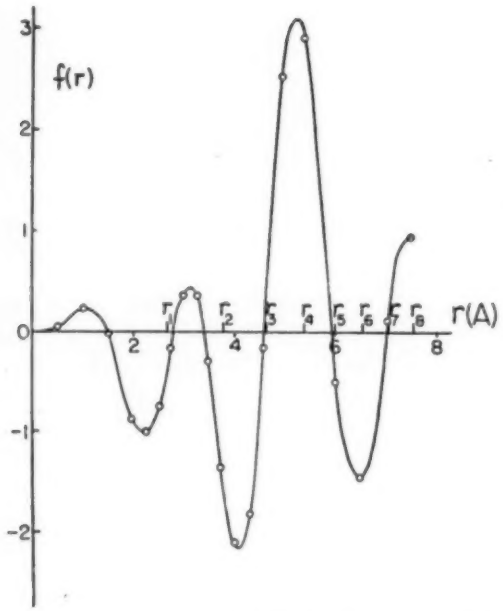


FIG. 5. Short-range order distribution function $f(r)$ for CuPt quenched from 930°C.

holder and in contact with the upper edge of the specimen.

Experimentally the measurements at temperatures above T_c included the short range order diffuse intensity, and also fluorescence radiation, Compton scattering, and two temperature dependent factors, nitrogen scattering and temperature diffuse scattering. By measuring the diffuse intensities at two temperatures below T_c and assuming a linear dependence of the change in intensity on the change in temperature, extrapolation determined the contributions of these latter factors at the temperatures above T_c where the short range order diffuse scattering was present.

Measurements were made over the range 5° to 25° in θ for temperatures of 25°C, 700°C, and 890°C in each run. The resultant diffuse intensities from three complete runs, normalized to electron units per atom, are plotted in Fig. 4b. Although there is appreciable scatter to the data, indicating the experimental difficulties encountered, the results in general show less intense, broader diffuse peaks than were obtained from the quenched samples, affirming the unreliability of quenched sample measurements of local order.

To perform the Fourier transform these intensity distributions were joined smoothly to the Laue Monotonic§ scattering above 25° in θ and extrapolated to obtain a smooth function below 5° in θ. Using a mechanical computer and tabulated values of sin Kx, the transform was carried through for 20 values of r between 0 and 7.5A for each of the high temperature runs and for the mean of the quenched sample distribution. A plot of the resulting $f(r)$ *vs* r for the mean of the quenched samples is shown in Fig. 5. The several modulations in the function at small r are attributed to errors introduced by joining the intensity distribution to the Laue Monotonic above 25° in θ. With the exception of these spurious modulations, this function $f(r)$ behaves quite regularly with respect to the radii of shells of atoms about some origin atom. It is zero for all odd shells and shows alternate negative and positive maxima for the even shells. Variations of 10 percent in the diffuse peak intensities or in the magnitude of the Laue Monotonic function caused variations in the amplitudes of the oscillations of the Fourier transforms, but the positions of the maxima and the intercepts remained essentially unchanged. The same general transform was obtained from the high temperature runs, except that the amplitudes of the oscillations were smaller and the intercepts beyond 6.5A showed some scatter.

The interpretation of these transforms in terms of the parameters $\bar{\alpha}_i$ requires that for i odd, $\bar{\alpha}_i=0$, while for i even, $\bar{\alpha}_i$ varies alternately negative and positive. From measurements of areas under the peaks of the transform, values of $\bar{\alpha}_2$, $\bar{\alpha}_4$, and $\bar{\alpha}_6$ were determined. The values from the high temperature runs, though showing scatter, are considerably lower than the corresponding parameters from the quenched specimens. These values of the parameters and estimates of their accuracy are recorded in Table II.

The physical meaning of the experimentally determined $\bar{\alpha}_i$'s is found by examining the ordered structure of CuPt. As is shown in Table II, the description of the ordered structure in terms of $\bar{\alpha}_i$'s requires for i odd, $\bar{\alpha}_i=0$, while for i even, $\bar{\alpha}_i$ varied alternately -1 and $+1$. This differs from the experimental $\bar{\alpha}_i$'s for equilibrium at 890°C only in the magnitude of the $\bar{\alpha}_i$'s. One concludes that above the critical temperature in CuPt, the atoms on the average manage to maintain local layering configurations, partially segregating onto alternate (111) planes, but with this tendency rapidly disappearing a few shells removed from any origin atom. There is not an "A-likes-B" type of short range order, with its attendant spherically symmetric atomic distribution, such as has been found in Cu_3Au,[4] but instead there is a local layering type of order.

The physical situation in CuPt above T_c may perhaps be imagined as the formation of a multitude of tiny anti-phase domains, each containing a reasonable amount of "long range order" in itself. If we take as an example domains the shape and size of one 32-atom ordered unit cell and attribute to each of these regions a value of $S=0.75$, approximately that found just below T_c, then by considering the number of ways these domains can bound one another, we can calculate the degree of short range order arising from this configuration. The result is

$$\bar{\alpha}_2=-.32,$$

§ Hönl corrections: $\Delta f_{Cu}=-2.5$; $\Delta f_{PtK}=-1.1$; $\Delta f_{PtL}=-6.1$. The first two values were calculated from Appendix III, James, the third estimated from measurements for tungsten recorded in R. W. James, *The Optical Principles of the Diffraction of X Rays* (G. Bell and Sons, Ltd., London, 1948), p. 187.

[4] J. M. Cowley, J. Appl. Phys. **21**, 24 (1950).

a value 60 percent larger than that experimentally determined. Though this particular model must offer only a rough approximation to a true physical domain description, it is still valuable in indicating that if one utilizes such a concept, one must expect either domains of this order of magnitude or else larger domains with a smaller degree of "long range order."

DISCUSSION

Though the determination of S *vs* T is thought inaccurate to 5 percent, the extrapolations and approximations made in determining the $\bar{\alpha}_i$ render their values rather inaccurate. Single crystal measurements following the technique of Cowley would probably afford more accurate determinations of the $\bar{\alpha}_i$, but facilities for the growth of such a crystal were not available.

It is of interest to consider the predictions of existing thermodynamic and statistical mechanical theories with respect to CuPt. All these theories based on the reasonable concept of short range interactions have assumed, among other things, that the interaction terms are spherically symmetric. Since the ordered structure of CuPt does not exhibit such symmetry, one might expect these theories to predict instability, as indeed is the case. Modification of A. H. Wilson's[5] generalization of Bethe's theory, Yang's[6] generalization of the Quasichemical method, and Zernike's[7] method to include propagation of order all predict $S=0$; and a modification of Cowley's[8] method predicts $S=0$ unless second neighbor forces are considered. One simple modification which permits structural definition is to consider the possibility of directional ordering forces and related parameters. Inclusion of this possibility allows both the Wilson and Cowley theories to predict stability, but the predicted S *vs* T dependence is of the continuous β-brass type, in contrast to the curve obtained experimentally.

The structure of the ordered form and the demonstration of local layer configurations above T_c give ample evidence that the stable structure of CuPt is that formed by segregation of the Cu and Pt atoms onto alternate planes, although there is no increase in the number of unlike neighbors as a result of such ordering and hence no contribution to the free energy of the system from the usual A-likes-B type of energy term. Above T_c the x-ray pattern is that of a cubic structure, but since the diffuse intensity indicates local layering, the cubic symmetry must be statistical only. While each little region of local layering has a particular layering direction, in different regions this layering direction can take with equal probability the four cube diagonal directions, thus giving statistical cubic symmetry. At low temperatures there is one layering direction, with the atoms perfectly segregated. With rising temperatures more and more atoms are located on the wrong layers, the long range order decreasing to about $S=0.77$ at the critical temperature. Above T_c the layering direction has four possible orientations, and one might consider the passage through the critical temperature as a disordering of the layering direction. If the regions of local layering are sufficiently small, enough additional entropy might result from this disordering of the layering direction to make this picture thermodynamically reasonable,

The author wishes to extend his thanks to Dr. B. E. Warren, both for the introduction to this topic and for his continued and helpful counsel through the course of this work.

TABLE II. Short range order parameters for CuPt from measurements of the diffuse scattering.

	Perfect order			Experimental $\bar{\alpha}_i$	
i	Like neighbors	Unlike neighbors	$\bar{\alpha}_i$ max	Quenched from 930°C	Held at 890°C
1	6	6	0	0.00	0.00
2	0	6	−1	−0.28	−0.20±0.06
3	12	12	0	0.00	0.00
4	12	0	1	−0.23	+0.12±0.05
5	12	12	0	0.00	0.00
6	0	8	−1	−0.14	−0.06±0.03

[5] A. H. Wilson, Proc. Cambridge Phil. Soc. **34**, 81 (1938).
[6] C. N. Yang, J. Chem. Phys. **13**, 66 (1945).
[7] F. Zernike, Physica **7**, 565 (1940).
[8] J. M. Cowley, Phys. Rev. **77**, 669 (1950).

A Note on Critical Reflections of Elastic Waves at Free Surfaces*

J. N. GOODIER AND R. E. D. BISHOP
Stanford University, Stanford, California
(Received September 17, 1951)

The wave motion in an elastic solid given by the theory of reflection at a free surface is evanescent for certain conditions—a P (plane irrotational) wave at grazing incidence and also an SV (plane equivoluminal) wave at grazing incidence. It is shown that by the application of suitable limiting processes wave motions can be obtained for these critical circumstances. The relation of these motions to the noncritical motions is analogous to the relation of critically damped to noncritically damped vibrations.

1. INCIDENT IRROTATIONAL WAVE

PLANE irrotational waves (P waves) traveling from left to right in Fig. 1 in the direction P (on the left) give rise, on reflection at a free surface $y=0$, to a reflected P wave (on the right) and also a reflected equivoluminal (SV) wave. The combined incident and reflected waves can be represented by displacement components

$$u=(\partial\phi/\partial x)+(\partial\psi/\partial y),\quad v=(\partial\phi/\partial y)-(\partial\psi/\partial x),\quad (1)$$

with

$$\phi=A_1\exp[ik(y\tan e+x-ct)]+B_1\exp[ik(-y\tan e+x-ct)],\quad (2)$$

$$\psi=B_2\exp[ik(-y\tan f+x-ct)];\qquad c=c_1/\cos e=c_2/\cos f,\quad (3)$$

these being solutions of the two-dimensional wave equations

$$c_1^2\nabla^2\phi=\partial^2\phi/\partial t^2,\quad c_2^2\nabla^2\psi=\partial^2\psi/\partial t^2,\quad (4)$$

where c_1 and c_2 are the two wave velocities in an infinite medium. The function ϕ gives irrotational motion, the function ψ equivoluminal motion. In (2) the A_1 term represents the incident P wave, the B_1 term the reflected P wave, and (3) represents the SV wave generated by reflection. The boundary conditions—that the normal stress and shear stress vanish at $y=0$—require that the constants A_1, B_1, B_2 be related according to[1]

$$B_1/[4\tan e\tan f-(1+3\tan^2 e)^2]=-B_2/[4\tan e(1+3\tan^2 e)]=A_1/[4\tan e\tan f+(1+3\tan^2 e)^2].\quad (5)$$

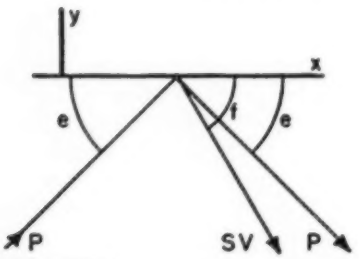


FIG. 1. Reflection of a P (plane irrotational) wave at a free surface.

Here Poisson's ratio has been given the value $\frac{1}{4}$ to simplify the formulas. With this value it follows that

$$c_1^2=3c_2^2\quad\text{and}\quad\cos^2 e=3\cos^2 f,\quad (6)$$

the latter stating that the intersection of a P-wave front with Ox travels along Ox at the same speed as the intersection of an SV wave front.

The motion vanishes altogether if the "angle of emergence" e is set equal to zero, for then (5) yields $B_1=-A_1$, $B_2=0$, and (2) and (3) show that ϕ and ψ vanish. Thus the functions (2) and (3) fail to give a motion in which there is a P wave traveling parallel to a free surface. Such a motion can be found however by allowing A_1 and B_1 to become infinite as e approaches zero, the product A_1e remaining finite, as in the formation of a doublet.

Expanding in powers of e we obtain from the second of (6)

$$\tan f=\sqrt{2}(1+\tfrac{3}{4}e^2+\cdots)$$

and from (5), dropping higher powers of e,

$$B_1=(-1+8\sqrt{2}e)A_1,\quad B_2=-4eA_1.$$

The functions (2) and (3) now become, with suitable development of the exponentials,

$$\phi=A_1e(8\sqrt{2}+2iky)\exp[ik(x-ct)]$$
$$\psi=-4A_1e\exp[ik(-y\sqrt{2}+x-ct)],\quad (7)$$

Thus the disappearance of the motion as e approaches zero is prevented by setting

$$\lim_{e\to 0}2A_1e=A_0.$$

Then (7) becomes, with c replaced by c_1,

$$\phi=A_0 4\sqrt{2}\exp ik(x-c_1t)+A_0iky\exp ik(x-c_1t)$$
$$\psi=-A_0\exp ik(-y\sqrt{2}+x-c_1t).\quad (8)$$

It is readily verified by substitution that the functions (8) satisfy the differential equations (4) and the free surface boundary conditions.

In (8) the first term in the expression for ϕ indicates an ordinary P wave traveling parallel to the free surface, the wave fronts being at right angles to the surface. It evidently corresponds to the incident P wave. The second term indicates an irrotational wave traveling in

* This note results from investigations carried out under a contract (N6onr 251 T.O. 12) between the ONR and Stanford University.

[1] See for instance K. E. Bullen, *An Introduction to Theoretical Seismology* (Cambridge University Press, London, 1947), arts. 6.4 and 6.5.

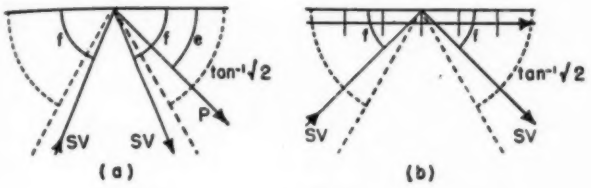


FIG. 2. Reflection of an SV (plane equivoluminal) wave at a free surface.

the same direction with the same speed, with fronts also at right angles to the surface. But the magnitude of the term varies over the wave front in proportion to y("Py" wave). The "reflected" SV wave expressed by the ψ-function in (8) is of the ordinary kind.

Irrotational waves with fronts at right angles to a free surface are encountered in impacts,[2] but their continuation into the interior must of course be quite different from that implied by Eqs. (8).

2. INCIDENT EQUIVOLUMINAL WAVE

The functions

$$\psi = A_2 \exp[ik(y\tan f + x - ct)] + B_2 \exp[ik(-y\tan f + x - ct)], \quad (9)$$
$$\phi = B_1 \exp[ik(-y\tan e + x - ct)],$$

represent an incident SV wave (A_2) with the reflected SV wave (B_2) and the reflected P wave (B_1) as in Fig. 2(a). The relation between e and f is given by the second of (6), which may be written as

$$3\tan^2 e = \tan^2 f - 2. \quad (10)$$

There is a real value of $\tan e$, and a reflected P wave as in Fig. 2(a), so long as $\tan f$ exceeds $\sqrt{2}$. When $\tan f$ is less than $\sqrt{2}$, $\tan e$ is imaginary, but the functions (9) still yield a real motion. The reflected P wave of Fig. 2(a) is replaced by a "surface wave" exponentially attenuated in the $-y$ direction, with fronts at right angles to the free surface. This is indicated by the succession of short vertical lines in Fig. 2(b).

The freedom of the boundary $y=0$ results in the following relations[1] between the constants in (9)

$$B_1/[4\tan f(1+3\tan^2 e)] = B_2/[4\tan e\tan f-(1+3\tan^2 e)^2] = A_2/[4\tan e\tan f+(1+3\tan^2 e)^2]. \quad (11)$$

No difficulty arises in the critical case $\tan f=\sqrt{2}$, $\tan e=0$. The motion disappears, however, as f goes to zero in Fig. 2(b), and again we apply a limiting process to obtain a result for this case.

Expanding in powers of f in (11), and using (10), we find

$$B_1 = -4fA_2, \quad B_2 = [-1+8(2/3)^{\frac{1}{2}}if]A_2, \quad (12)$$

dropping higher powers of f. Then, setting

$$\lim_{f\to 0} A_2 f = -1/2A_0'i,$$

we obtain from (9)

$$\psi = A_0' 4(2/3)^{\frac{1}{2}} \exp[ik(x-c_2t)] + A_0' ky \exp[ik(x-c_2t)], \quad (13)$$
$$\phi = A_0' 2i \exp[(2/3)^{\frac{1}{2}}ky]\exp[ik(x-c_2t)].$$

The first term of the ψ function represents an SV wave traveling parallel to the surface, the second an equivoluminal wave traveling similarly, but with variation over a wave front according to the factor y ("SVy" wave). The ϕ function is an irrotational surface wave of the ordinary kind, as indicated in Fig. 2(b).

The results expressed by Eqs. (8) and (13) are summarized in Figs. 3(a) and (d). Terms containing y as a factor are indicated as "Py" or "SVy" waves. Figure 3(b) is obtained from Fig. 3(a) by reversal and turning over. Figure 3(c) represents Eqs. (9) and (11) for the critical case $\tan f=\sqrt{2}$, $\tan e=0$. Figures 3(b) and 3(c) each represent reflection of an incident SV wave with $\tan f=\sqrt{2}$. Other types of reflection of the same incident wave can be formed by combination of these two. Evidently reflection at this angle is not a unique process.

3. *Py* AND *SVy* WAVES IN THE INFINITE MEDIUM

The Py waves appear as part of the most general irrotational plane wave system which can be propagated with velocity c_1. Under these conditions, dealing directly with displacements, we have (in two dimensions—plane strain)

$$\partial v/\partial x = \partial u/\partial y, \quad c_1^2\nabla u^2 = \partial^2 u/\partial t^2, \quad c_1^2\nabla v^2 = \partial^2 v/\partial t^2, \quad (14)$$

and

$$u = u(y, x-c_1t), \quad v = v(y, x-c_1t), \quad (15)$$

for propagation in the x direction. These forms of u and v satisfy the equations

$$c_1^2(\partial^2 u/\partial x^2) = \partial^2 u/\partial t^2, \quad c_1^2(\partial^2 v/\partial x^2) = \partial^2 v/\partial t^2,$$

and therefore if they are also to satisfy (14) it is necessary that

$$\partial^2 u/\partial y^2 = \partial^2 v/\partial y^2 = 0,$$

which reduce (15) to forms linear in y. Having regard

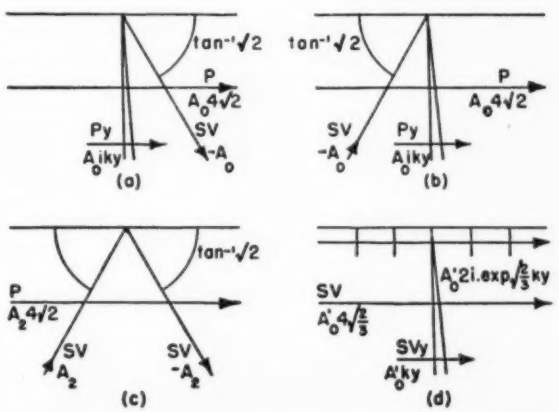


FIG. 3. Several wave systems consistent with a free surface.

[2] F. Sauter, Z. angew. Math. Mech. **30**, 149–153, 203–215 (1950).

to the irrotationality condition in (14) these forms are

$$u=f_1(x-c_1t)+y(\partial/\partial x)f_2(x-c_1t), \quad v=f_2(x-c_1t),$$

f_1 and f_2 denoting arbitrary functions. The f_1 system is the ordinary plane irrotational P wave. The f_2 system is of the Py type.

Similarly SVy waves appear as part of the most general equivoluminal plane wave system which can travel with velocity c_2. This general system is

$$u=f_3(x-c_2t), \quad v=f_4(x-c_2t)-y(\partial/\partial x)f_3(x-c_2t),$$

and the f_3 system is of the SVy type.

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 1 JANUARY, 1952

A Direct-Current Network Analyzer for Solving Wave-Equation Boundary-Value Problems

GEORGE W. SWENSON, JR.*
Washington University, St. Louis, Missouri

AND

THOMAS J. HIGGINS†
University of Wisconsin, Madison, Wisconsin

(Received August 24, 1951)

The analogy between the difference equations of a wave-equation boundary-value problem and the current-law equations of a direct-current network is developed. This theory permits construction of a dc network analyzer for solution of the wave equation. Some of the network parameters are necessarily negative. These "negative resistors" are simulated by active dc circuits. The determination of characteristic frequencies, characteristic functions, and response to simple harmonic forces is illustrated by solution of an irregularly shaped membrane. Methods are discussed for solving problems involving inhomogeneous or anisotropic media or various types of boundary conditions.

INTRODUCTION

THE problem of determining the characteristic frequencies and characteristic functions of a bounded medium whose performance is governed by the classical wave equation is, in general, an inherently difficult one. Analytical methods are known only for the eleven "Eisenhart" coordinate systems in which the wave equation is separable. Accordingly, unless the boundaries of the medium coincide with parametric surfaces of one of these coordinate systems, the problem may be said to be insoluble by analytical means.

Needless to say, many situations arise in practice in which it is necessary to obtain solutions of problems not encompassed in this limited group; for example, the determination of the cut-off frequencies of waveguides or the sound pressure distribution within an acoustical resonator. Accordingly, a technique of numerical analysis[1,2] has been developed, based on the calculus of finite differences, by which the characteristic functions and, with more difficulty, the characteristic frequencies can be found. Briefly, this method comprises dividing the medium into a number of elementary subareas or subvolumes, and then writing for each of these elements a difference equation which approximates the wave equation at that point. The partial differential equation is thereby transformed into a set of simultaneous, linear, algebraic equations, which can be solved to yield an approximate solution to the boundary-value problem. Unfortunately, when the number of difference equations is large, as it must be in order to yield an accurate approximation, the labor required to evaluate the unknown quantities is great. Nevertheless, as the most feasible method available, the technique has been widely employed.

The finite-difference technique has been applied even more widely to other partial differential equations, including Laplace's, Poisson's, and the diffusion equation. To avoid the labor involved in solving the corresponding difference equations, analog solution by electrical networks which have equations of voltage or current analogous to the difference equations has been utilized. Experimental solutions can be obtained thereby without appreciable computation.

Recently, successful applications of alternating current networks to the solution of wave equation problems have been reported[3,4,5]. The applications were limited to two-dimensional problems and three-dimensional problems with axial symmetry. Networks to simulate nonrestricted three-dimensional problems have

* Associate Professor of Electrical Engineering.
† Professor of Electrical Engineering.
[1] R. V. Southwell, *Relaxation Methods in Theoretical Physics* (Oxford University Press, London, 1946).
[2] L. Collatz, *Eigenwertprobleme und ihre numerische Behandlung* (Chelsea Publishing Company, New York, 1948).
[3] Spangenberg, Walters, and Schott, Proc. Inst. Radio Engrs. 37, 724, 866 (1949).
[4] Whinnery, Concordia, Ridgeway, and Kron, Proc. Inst. Radio Engrs. 32, 360 (1944).
[5] R. H. MacNeal, "Solution of partial differential equations by means of electrical networks," Ph.D. thesis, California Institute of Technology, Pasadena, 1949.

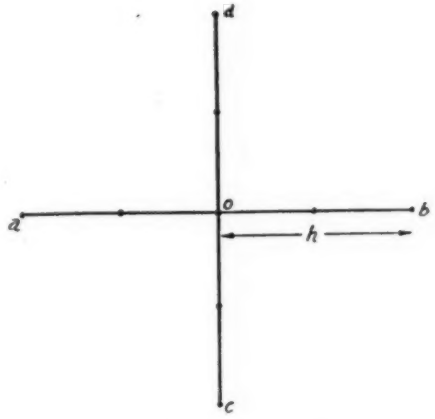


FIG. 1. Finite-difference grid.

been suggested, but these networks would be so complicated and expensive to build that they appear impractical for actual application. Recently, an account[6] has appeared of a direct-current network analyzer capable of solving the wave equation, but apparently no actual data relative to its application to the wave equation have been published to date.

This paper is concerned with a direct-current network analyzer specifically designed for two-dimensional problems but capable of extension to three dimensions, for solving the difference equations of a wave-equation boundary-value problem. The device as now constructed[7] can be used not only for determining characteristic functions and characteristic frequencies, but also the steady-state response to a sinusoidally time-varying disturbance. A direct-current network analyzer has the advantage of requiring as circuit elements only resistors, lower in cost and capable of more precise calibration than the inductors and capacitors of an ac network analyzer. The network is unique in that it makes use of "negative resistors," as required for satisfaction of the difference equations. These negative elements are manually adjusted, the result being an iterative process, in some respects similar to the iteration methods for numerical solution of systems of algebraic equations. The mechanical labor involved, however, is far less than the arithmetical labor of the numerical methods, so that a great saving in time is obtained.

BASIC THEORY

The basic theory and method of application of the dc analyzer are most easily developed in terms of the vibrating membrane. This problem is comparatively simple as regards boundary conditions and analogous to many technically important systems governed by the wave equation. Consider a membrane whose boundaries are fixed, and whose quiescent surface is plane. For small, simple-harmonic, transverse motions the displacements are of the form $z=\phi(x, y)e^{j\omega t}$ wherein $\phi(x, y)$ is given by the wave equation in Helmholtz form,

$$\nabla^2\phi+(\omega/c)^2\phi=0, \tag{1}$$

where ω is the angular frequency of a characteristic mode of vibration and c is the characteristic phase velocity as determined by the density and tension of the membrane. For the present, these latter two quantities are considered as constant, so that c is uniform throughout the membrane.

Let $\phi_o(x_o, y_o)$ designate the transverse displacement at point O of Fig. 1. Then, expanding $\nabla^2\phi(x, y)$ and $\phi(x, y)$ about point O in double Taylor series of uniform increment $\Delta x=\Delta y=h$, substituting in (1), and neglecting all terms of order greater than h^2 yields

$$-\phi_a-\phi_b-\phi_c-\phi_d+\beta\phi_o=0, \tag{2}$$

in which

$$\beta=4-\omega^2h^2/c^2. \tag{3}$$

An equation of the form of (2) applies at each interior point in the finite-difference grid. If β is known, the displacements at all points of the grid can be determined by solving the set of simultaneous equations thus obtained. It is to be noted that a nontrivial solution of this set of homogeneous equations exists only if β is one of the characteristic values. A method of finding at least some of these values is discussed below.

Figure 2 illustrates a rectangular membrane divided by a grid with nine interior points. From (2) and (3), the difference equations corresponding to points 1, 2, and 5, are, respectively,

$$(4-\omega^2h^2/c^2)\phi_1-\phi_2-\phi_4=0, \tag{4}$$

$$-\phi_1+(4-\omega^2h^2/c^2)\phi_2-\phi_3-\phi_5=0, \tag{5}$$

$$-\phi_2-\phi_4+(4-\omega^2h^2/c^2)\phi_5-\phi_6-\phi_8=0. \tag{6}$$

Figure 3 displays the network analog to these difference equations. If V_n is the node-to-ground voltage at node n, this node being positive with respect to ground, the Kirchhoff's current-law equations for nodes 1, 2, and 5

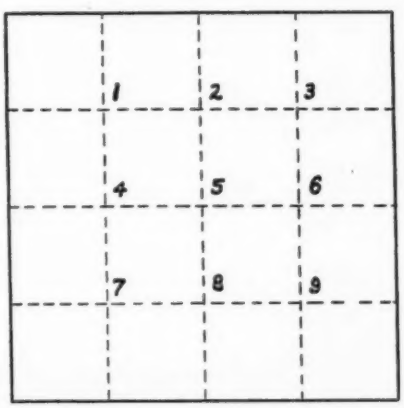


FIG. 2. Rectangular membrane with finite-difference grid.

[6] G. Liebmann, Brit. J. Appl. Phys. 1, 92 (1950).
[7] G. W. Swenson, Jr., "A direct current network analyzer for solving wave-equation boundary-value problems," Ph.D. thesis, University of Wisconsin, Madison, 1951.

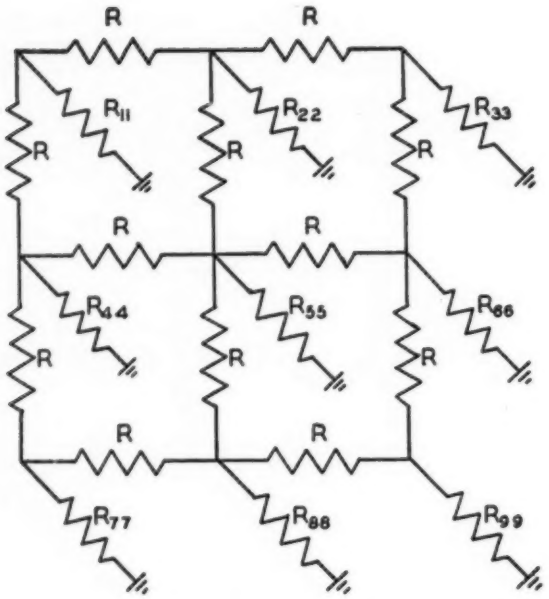


FIG. 3. Network analog to the problem of Fig. 2.

are, respectively,

$$(2+R/R_{11})V_1-V_2-V_4=0, \quad (7)$$

$$-V_1+(3+R/R_{22})V_2-V_3-V_5=0, \quad (8)$$

$$-V_2-V_4+(4+R/R_{55})V_5-V_6-V_8=0. \quad (9)$$

Comparison of (4), (5), and (6), respectively, with (7), (8), and (9), gives each ϕ_n as equal to its corresponding V_n, provided that

$$R_{11}=R/(2-\omega^2h^2/c^2), \quad (10)$$

$$R_{22}=R/(1-\omega^2h^2/c^2), \quad (11)$$

$$R_{55}=-Rc^2/(\omega^2h^2). \quad (12)$$

Nodes 1 and 2 correspond to membrane points *adjacent* to the boundaries of the membrane. Respectively, they have two and three neighboring nodes whose voltages to ground are not zero. For such nodes the values of R_{nn} given by (10) and (11) are positive for h sufficiently small to yield acceptable accuracy. At a node surrounded by four nonzero neighbors, however, (12) shows that the node-to-ground resistance must be negative.

The necessary negative resistor can be synthesized from positive resistors and a voltage source (Fig. 4). Resistor R_n is adjusted on a Wheatstone bridge to a value equal to the magnitude of the desired negative resistance, R_{nn}. If, now, R_s is adjusted manually so that voltmeter V reads zero, a current V_n/R_n exists in R_n in the direction toward node n. Accordingly, when connected to the analyzer, this circuit comprises the desired negative resistor.

In solving a problem a number of these negative resistors are required. Because adjustment of each influences the adjustment of all the others, an iterative process of adjustment is required. The desired solution is attained when all voltages V are zero. The process converges over an important range of values of β, as described below.

CONSTRUCTION AND OPERATION OF THE ANALYZER

An analyzer capable of solving 100 simultaneous equations has been constructed. The resistors R_s and R_n are inexpensive, two-watt, wire-wound potentiometers, equipped with banana plugs for ease of adjustment and replacement. Resistors R and R' are of one-percent tolerance, are 118 ohms and 236 ohms, respectively, and are permanently wired. For constancy of voltage, storage batteries are used for excitation of the network. One battery of thirty volts and one of six volts are used, the terminals of all cells being made available for voltage adjustment. An electronic dc millivoltmeter capable of indicating one millivolt is used as the null indicator for the negative resistors. A more precise, high-resistance voltmeter is used for obtaining the values of the solution.

To set up the analyzer for a problem the desired value of ω^2h^2/c^2 is chosen and all the node-to-ground resistors (positive and negative) are adjusted on a Wheatstone bridge in accordance with (10), (11), and (12). Then each potentiometer R_s is adjusted (in turn) until its corresponding voltage, V, becomes equal to zero. This procedure is repeated until all the V's in the network are zero simultaneously.

If the network is set up as just described, then because the negative resistors are the only energy sources in the network, the trivial solution of all V_n's equal to zero is obtained. This solution corresponds to zero displacement of the membrane. The remedy is to apply a battery between some node of the network and ground. This constraint simulates a fixed displacement, ϕ_n, of the corresponding point of the membrane, and forces the network to a solution which conforms to the established value at the forced point. This battery also serves the purpose of fixing the order of magnitude of the node-to-ground voltages, V_n, representing the solution.

If the frequency ω used in setting up the network is not a characteristic frequency of the system, a current will exist in the battery, the value of this current is proportional to the force necessary to produce the displace-

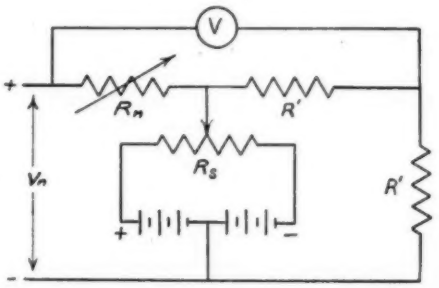


FIG. 4. The negative resistor.

ment ϕ_n. Thus, in general, problems solved by the analyzer are those of forced vibration.

DETERMINATION OF CHARACTERISTIC FREQUENCIES

Up to this point it has been assumed that characteristic values of ω^2h^2/c^2 were available for use in setting up the analyzer. In an actual problem, however, it may be that determination of these values is one of the purposes of the investigation. To this end it is to be noted that at a characteristic frequency no external force is required to produce a finite displacement of the membrane. Because current supplied by the battery used to obtain a nontrivial solution is analogous to externally applied force, a null value of this current indicates "resonance." Using an arbitrary value of ω^2h^2/c^2, the analyzer is set up and adjusted to obtain a solution. When all the negative resistors are properly balanced, the battery current is measured and plotted as a function of frequency. By repeating the process, sufficient points on the graph are obtained to enable determination of the characteristic frequency by interpolating to the point at which the current is zero. An example of this procedure is advanced below in connection with the illustrative problem of the truncated square membrane. This method has been used to check the calculated value of the lowest natural frequency of a square membrane, using 100 difference equations, to within 0.5 percent. With more accurate analyzer components and greater care in adjustment, the accuracy of this technique could undoubtedly be increased considerably.

Experience has shown that the adjustment process for the negative resistors is convergent for frequencies from zero through the first characteristic value and up to a point just below the first "antiresonant frequency." At this latter value an infinite force would be required to produce a finite displacement. No schedule of adjustment has been found which will secure convergence at higher frequencies. However, certain higher modes, in which the nodal lines are lines of symmetry, can be investigated piecemeal. It can be shown with the aid of Rayleigh's principle that if a membrane possesses a line of symmetry, a characteristic mode of vibration exists in which a nodal line coincides with the line of symmetry. Thus, for such a mode, the characteristic frequency can be ascertained by determining the lowest frequency of one of the *zones* of the membrane bounded by nodal lines or fixed boundaries. Although the analyzer does not give directly the solution to an assymmetrical higher mode, it is of great assistance in connection with conventional numerical techniques involving Rayleigh's principle[8] and orthogonalization.

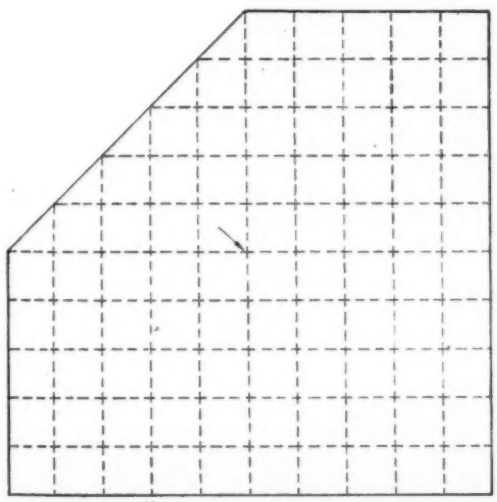


FIG. 5. A truncated square membrane.

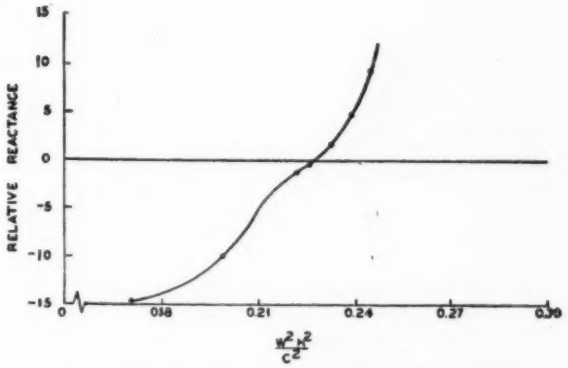


FIG. 6. Determination of characteristic frequency of truncated square membrane.

FORCED VIBRATION

Problems in forced vibration, very difficult of analysis by analytical methods, are solved with ease on the network analyzer. The equation for steady-state response to a simple harmonic force is

$$\nabla^2\phi+(\omega^2/c^2)\phi=P/T, \tag{13}$$

where P is the applied transverse force over area h^2 at point O and T is the tension (force per unit length) of the membrane. The typical difference equation approximating (13) at point O is

$$-\phi_a-\phi_b-\phi_c-\phi_d+\beta\phi_o=-P_o/T. \tag{14}$$

An extension of the previous discussion, which led to (10), (11), and (12), shows that the current which must be inserted into node n to simulate the force at that point is

$$i_n=-P_n/(RT). \tag{15}$$

These currents are conveniently derived from a common high-voltage supply through dropping resistors large compared with R. From knowledge of the current input to a node, the frequency, and the node-to-ground voltage, the mechanical impedance of the corresponding point of the membrane can be found.

[8] G. Temple and W. G. Bickley, *Rayleigh's Principle* (Oxford University Press, London, 1933).

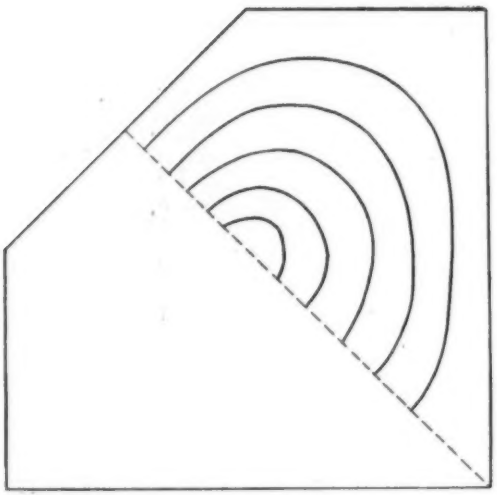


FIG. 7. Vibration of truncated membrane: $\omega^2h^2/c^2=0.170$.

ILLUSTRATIVE PROBLEM: TRUNCATED SQUARE MEMBRANE

Figure 5 illustrates a membrane of irregular shape, divided by a finite-difference grid. The analyzer setup involves 71 node points, including 43 which require negative resistors. By the method described under "Determination of Characteristic Frequencies," the lowest characteristic value of ω^2h^2/c^2 was found to be 0.227 (Fig. 6). Figure 7 shows the form of the response of the membrane to a force whose frequency is below the first characteristic frequency; Fig. 8 shows the characteristic function for the first mode; and Figs. 9, 10, and 11 show the form of response to forces of frequency higher than that of the lowest mode. In each case, relative displacements are indicated by means of contour lines of equal displacement, the contour interval being uniform for a given figure. The point of application of force is indicated by the arrow of Fig. 5.

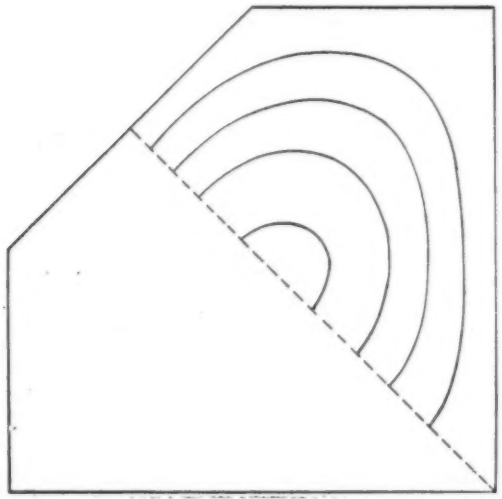


FIG. 8. Lowest characteristic function of truncated membrane: $\omega^2h^2/c^2=0.227$.

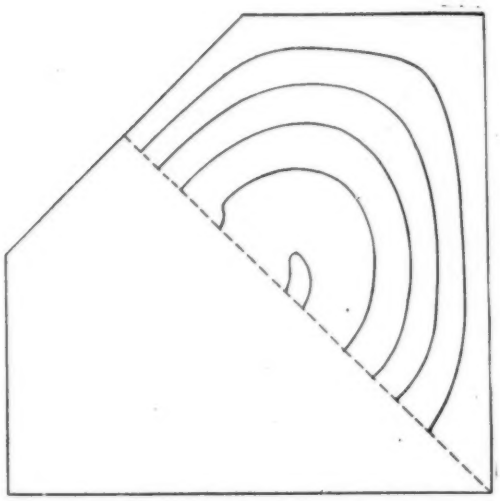


FIG. 9. Vibration of truncated membrane: $\omega^2h^2/c^2=0.245$.

INHOMOGENEITY AND ANISOTROPY

A vibrating membrane whose characteristic (unbounded) velocity of phase propagation, c, varies from point to point within the boundaries can be investigated quite easily by means of the analyzer. The change in velocity results in the values of node-to-ground resistors varying from node-to-node in the network. These are calculated from (10), (11), and (12), using the value of c at the point of the membrane in question. Characteristic frequencies, characteristic functions, and response to an applied force are determined in the same manner as for a homogeneous membrane.

The solution of the wave equation in an anisotropic medium can often be accomplished on the analyzer by first making such a transformation of variables that the velocity of propagation at any point is independent of direction. For example, for a membrane whose superficial density is uniform but whose x and y components of velocity are T_1 and T_2, respectively, the wave equa-

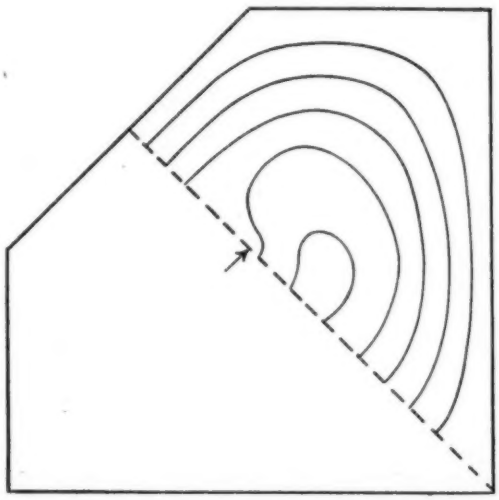


FIG. 10. Vibration of truncated membrane: $\omega^2h^2/c^2=0.259$.

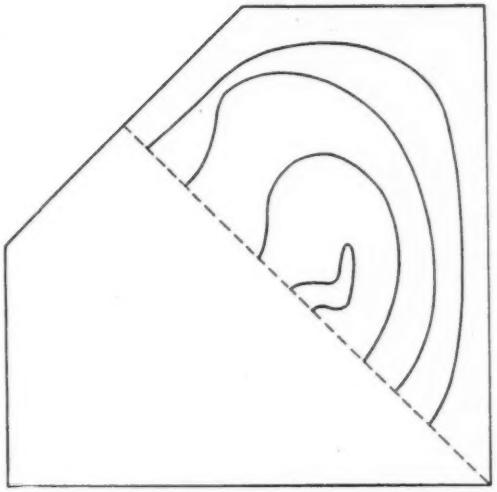


FIG. 11. Vibration of truncated membrane: $\omega^2 h^2/c^2 = 0.290$.

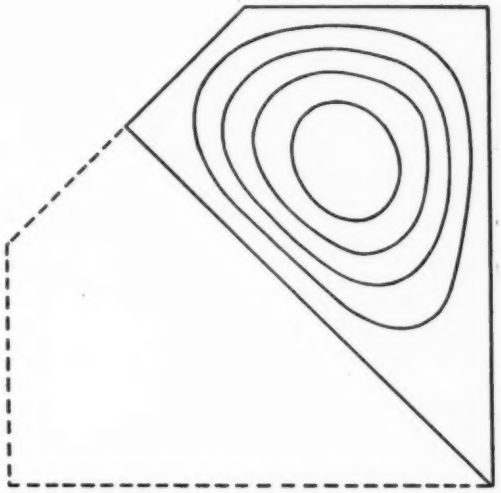


FIG. 12. Higher mode of truncated membrane: $\omega^2 h^2/c^2 = 0.501$.

tion can be put into appropriate form for the analyzer by using the transformations $\theta = x/T_1^{\frac{1}{2}}$, $\psi = y/T_2^{\frac{1}{2}}$.

BOUNDARY CONDITIONS

Although discussion up to this point is restricted to problems in which the boundary condition is that of zero displacement, other boundary conditions can also be accommodated on the analyzer. One of these is the case in which the normal derivative of ϕ is zero on the boundary. This condition demands that an extra row of network nodes be placed adjacent to, but outside of, the boundary nodes. The zero-derivative condition is established by making equal to one another the node-to-ground voltages of the two opposite nodes, one inside and one outside of the boundary, and both adjacent to it. The nodes outside the boundary need not have the conventional node-to-ground resistances, but merely adjustable voltage sources. The differences between voltages of the opposite nodes are adjusted to zero by means of these sources during the regular adjustment of the negative resistors in the rest of the network. This technique also permits a more detailed solution of symmetrical modes than would otherwise be possible in a network of a given number of nodes. Any equal-phase portion of the membrane enclosed by fixed boundaries and lines of symmetry can be analyzed separately, using the full extent of the available analyzer, the boundary condition along a line of symmetry being that the normal derivative is zero.

Boundaries on which the normal derivative is a nonzero constant or a known function of position can be simulated by an extension of the same method. It is merely necessary that the difference between two opposite nodes adjacent to the boundary be of the correct nonzero value. This can conveniently be achieved during the regular adjustment process by the use of voltage-dividing resistors.

ACKNOWLEDGMENT

The authors wish to express their appreciation to Professor Harold A. Peterson for his aid and encouragement, and to the University of Wisconsin Engineering Experiment Station for financial aid during part of the project discussed here.

Phenomena Associated with the Flight of Ultra-Speed Pellets.* Part I. Ballistics

WILLIAM A. ALLEN, JOHN S. RINEHART, AND W. C. WHITE
Michelson Laboratory, U. S. Naval Ordnance Test Station, Inyokern, China Lake, California
(Received July 16, 1951)

Small metallic pellets have been fired at velocities up to 6 km/sec by means of modified shaped charges. Previous work in this field has been confined largely to observations of meteors and conventional shaped charges. Data obtained on ballistic phenomena verify the conventional drag force determinations extrapolated to large Mach numbers. Steel pellets at velocities of 2.5 km/sec lose little mass in flight. Pellets composed of lighter metals, such as aluminum at a velocity of 5 km/sec, burn vigorously with the emission of much light and heat. It has been found that the motion of both nonburning and burning pellets can be represented within experimental error by the empirical equation

$$v=v_0e^{-\alpha s},$$

where v_0 is the initial velocity of the pellet, v is its velocity after it has traveled a distance s through the air, and α is a constant.

INTRODUCTION

APPLIED to macroscopic phenomena, the term "ultra-speed" is defined as being generally applicable to velocities in excess of those usually encountered in ordinary ballistics and specifically to a velocity spectrum of 2 to 70 km/sec. Pellets that have velocities in the range from 2 to 6 km/sec can be achieved under laboratory conditions by means of shaped charges.[1] Meteors whose velocities lie in the range from 8 to 68 km/sec have been studied extensively by astronomers. Excellent reviews of the work in this field have been published.[2,3]

The purpose of this report is to present data obtained recently on the flight of individual metallic pellets that weigh from a fraction of a gram to several grams. Pellet velocities have ranged from 2.5 to 6 km/sec. The ballistics of the pellets are described in Part I of this report; the spectroscopic character of the luminous trails associated with the pellets, in Part II; and the general features of luminosity, in Part III. For convenience, the units were taken as meters, grams, and milliseconds.

GENERAL BEHAVIOR OF PELLETS

A modified shaped charge has been used to impel $\frac{1}{2}$-in. diameter pellets of volume $\frac{1}{4}$ cc. Pellets of many different metals have been fired. Those made from soft metals such as zinc, cadmium, and tin did not hold together and were projected as a spray of small fragments or droplets. Such pellets were not used in the present ballistic and spectroscopic studies. The materials found most suitable were steel, aluminum, magnesium, titanium, and a magnesium-rich magnesium-lithium-aluminum alloy.[4] The densities and corresponding weights of the several kinds of pellets are listed in Table I.

TABLE I. Densities and weights of pellets.

Pellet material	Density (g/cc)	Weight (g)
Steel	7.70	1.95
2S—O aluminum	2.69	0.68
Magnesium	1.74	0.44
Mg—Li alloy	1.44	0.36

Steel pellets deform considerably under the impact of the high explosive. Little mass is lost during flight through the air. The steel pellet is faintly luminous during flight, and it is believed that this light is due to the luminosity of the associated shock wave. No spectrograms were obtained from the light caused by the steel pellets.

Many of the pellets made of aluminum, titanium, and the special magnesium-lithium alloy did not break up during their flight through the air. No pellet of any of these materials, however, was ever recovered intact. Each of these pellets produced an intensely luminous trail. A photograph of a burning magnesium-lithium pellet is reproduced in Fig. 1. Many excellent spectrograms have been obtained of the trails and will be described in Part II of this report.

The magnesium pellets usually broke up in flight. Nevertheless, several luminous trails have been photographed, and some spectrograms obtained.

THE EXPERIMENTAL ARRANGEMENT

Most of the pellets were projected horizontally; a few were projected vertically. Many of the pellets were recovered in a thick Celotex target usually placed about 30 ft from the charge. Penetration into the Celotex was used as a rough measure of the velocity of the pellet. For more precise work, the velocities were determined from photographs taken with a General Radio Type

* Work supported by Armament Branch, ONR.
[1] Birkhoff, MacDougall, Pugh, and Taylor, J. Appl. Phys. **19**, 563 (1948).
[2] F. L. Whipple, Revs. Modern Phys. **15**, 246 (1943).
[3] Lovell, Prentice, Porter, Pearse, and Herlofson, Repts. on Prog. in Phys. **XI**, 389 (1946–47).
[4] This is a special alloy developed by the Battelle Research Foundation. Its composition is 80.6 percent magnesium, 13.4 percent lithium, and 6 percent aluminum. The material used here was supplied by the Naval Research Laboratory.

561-AG High-Speed Camera.[5] The photographic record proved to be sufficiently accurate for purposes of drag determination. A typical streak photograph of an aluminum pellet fired over a 30-ft trajectory into a Celotex target is reproduced in Fig. 2.

Other experimental techniques have been utilized. Inconclusive experiments were made with ballistic pendulums for the purpose of measuring possible mass-loss; experiments were made using x-ray film yaw cards in an attempt to measure the presentation area[6] of the pellets. In particular, a microflash technique was used[7] from which it was possible to obtain measurements on both luminous and nonluminous pellets. With this technique measurements can be made upon the products of combustion, or the smoke trail, which persists for some time along the trail of burning pellets. Figure 3 shows a burning magnesium pellet apparently bursting into several pieces at a point about 10 ft along its trajectory. The discrete frames, numbered in sequence 1, 2, ···, following the initial luminous phenomena, were obtained at a flashing rate of 2000 frames per second.

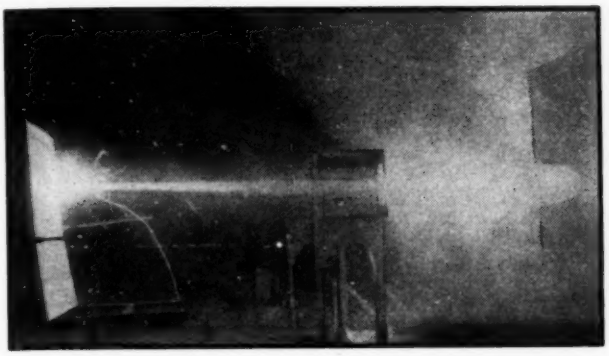


FIG. 1. The luminous trail generated by a 6 m/msec magnesium-lithium pellet moving from right to left. The ballistic data for this round are plotted in Fig. 4 (round 2). The box-like object in the center of the field is a specially constructed field spectrograph.

EQUATIONS OF MOTION

Streak photographs of a large number of pellets have been measured to obtain an analytical expression that describes the motion of the pellet. It has been found that the motion of both nonburning and burning pellets can be represented within experimental error by the empirical equation

$$v = v_0 e^{-\alpha s}, \tag{1}$$

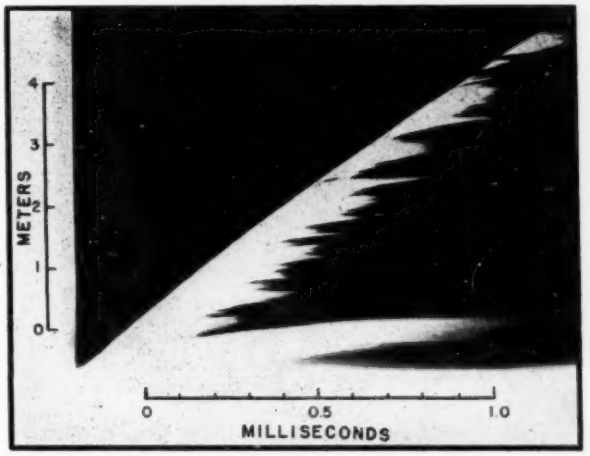


FIG. 2. Streak photograph of an aluminum pellet.

where v_0 is the initial velocity of the pellet, v is its velocity after it has traveled a distance s through the air, and α is a constant whose value is affected by several factors, among which are the pellet material and the number of pieces into which the pellet breaks up.

Values of v_0 and α for the several pellets were obtained from streak photographs of the type reproduced in Fig. 2. In such photographs the time axis, t, runs parallel to the film, and the space axis, s, lies in the film perpendicular to the time axis. Space coordinates of the trajectory were measured to the nearest micron on the original 35-mm film at increments of one-half mm along the time axis. Velocities at points along the trajectory were obtained by differencing appropriate space coordinates. The leading edge of the trace in Fig. 2 is somewhat irregular. This irregularity, which seems to be due to the yawing of the pellet, is responsible for much of the observed experimental scatter. Velocities of four magnesium-lithium pellets determined in this

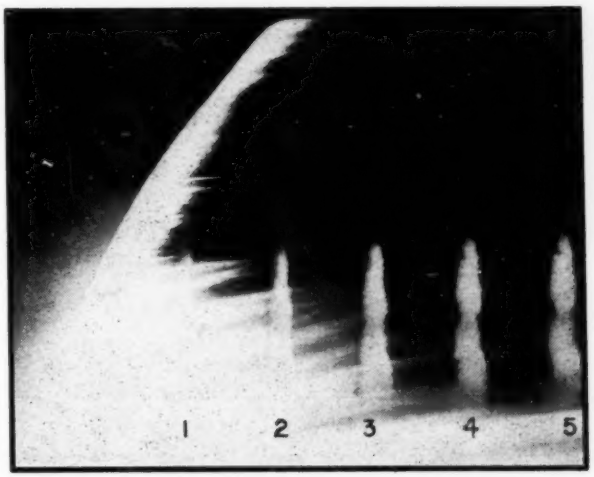


FIG. 3. A burning magnesium pellet. The discrete frames, numbered in sequence 1, 2, ···, following the initial luminous phenomena, were taken at a framing rate of 2000 per second.

[5] This camera was intended for taking high speed oscillograms at film velocities from about 35 to 100 ft/sec. The camera was modified for the present application in order to obtain the required magnification.

[6] The area presented by an object in the direction of the trajectory is defined as the area of the orthogonal projection of the object on the plane normal to the flight line. It is the area of the shadow of the object, or the area of the silhouette of the object as observed from a great distance along the trajectory.

[7] J. S. Rinehart, Rev. Sci. Instr. **21**, 939 (1950).

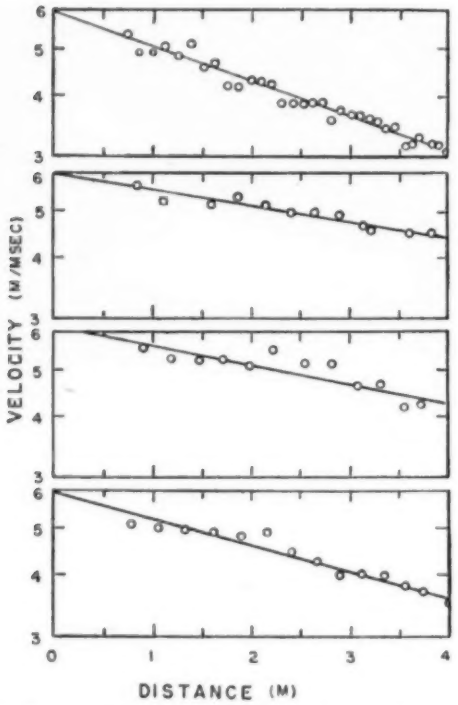


FIG. 4. Experimental data for four magnesium-lithium pellets. The straight lines represent the equation, $v=v_0e^{-\alpha s}$. The rounds are numbered 1, 2, 3, and 4 from top to bottom. The respective values of v_0 are 5.91, 5.96, 6.13, 5.92, and the respective values of α are 0.164, 0.076, 0.090, and 0.124.

manner are plotted in Fig. 4. The plots have been made on semi-log paper and a straight line fitted to the points by the method of least squares. The value of α is obtained from the slope of the line, and the value of v_0 is obtained from the intercept. Experimentally determined values of v_0 and α for the several kinds of pellets fired are listed in Table II.

It follows from Eq. (1) that the acceleration, a, of the pellet will be given by the relation

$$a=-\alpha v^2. \tag{2}$$

The motion can be expressed in terms of time by the relations

$$s=\alpha^{-1}\log\beta \tag{3}$$

$$\frac{ds}{dt}\equiv v=v_0\beta^{-1} \tag{4}$$

$$\frac{d^2s}{dt^2}\equiv\frac{dv}{dt}\equiv a=-\alpha v_0^2\beta^{-2}, \tag{5}$$

TABLE II. Experimentally determined constants for equations of motion.

Pellet material	Density (g/cc)	v_0 (m/msec)	α (m^{-1})
Steel	7.70	2.5	0.018
2S−O aluminum	2.70	5.0	0.050
Magnesium	1.67	5.8	0.144
Mg−Li alloy	1.44	6.0	0.076

where

$$\beta=1+\alpha v_0 t. \tag{6}$$

Plots of Eqs. (3), (4), and (5) for each of the several kinds of pellets are reproduced in Fig. 5. The values of v_0 and α listed in Table II were used in making the plots.

Tremendous forces act on ultra-speed pellets. The magnitude of the pressure caused by the high explosive probably is of the order of 100,000 atmospheres.[8,9] The force of the air opposing the motion of an ultra-speed pellet can be calculated. In the case of the magnesium-lithium pellet, the maximum deceleration found from Eq. (2) is about 294,000 times the acceleration due to gravity. If the area upon which the force is acting is

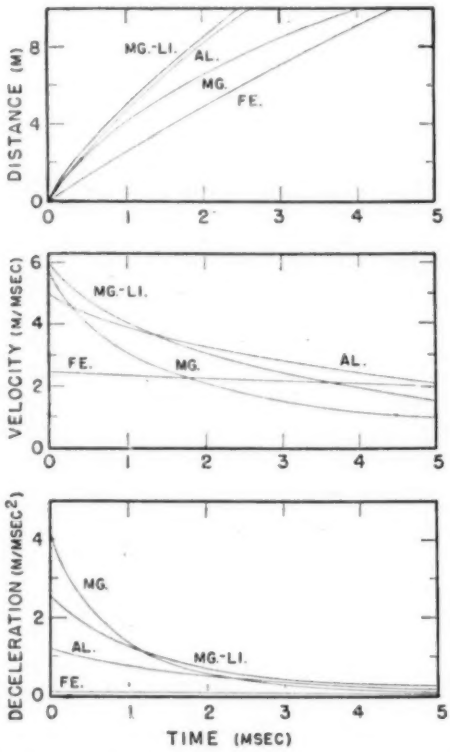


FIG. 5. Distance, velocity, and acceleration as functions of time for pellets composed of different kinds of materials.

assumed to be 0.68 cm², then the pellet is subjected to a pressure of about 150 atmospheres. The classical aerodynamic calculation predicts a pressure rise of 372 atmospheres for the stagnation pressure behind a detached shock wave travelling at 6.0 m/msec.

NONBURNING PELLETS

It is to be expected from conventional aerodynamic considerations that the motion of a pellet that neither changes shape nor loses mass during flight is given by

[8] R. H. Kent, J. Appl. Phys. **13**, 348 (1942).
[9] C. S. Robinson, *Explosions, Their Anatomy and Destructiveness* (McGraw-Hill Book Company, Inc., New York, 1944).

Eqs. (1) and (2). Eq. (2) is usually written in the form

$$\text{Drag force} = -K_D \rho d^2 v^2, \tag{7}$$

where K_D is the drag coefficient, ρ is the density of the air, d is the diameter of the projectile, and v is the velocity of the projectile. It is of interest to compare the drag on the steel pellets fired here with drag data obtained on spheres by Charters and Thomas[10] and on cubes by Rinehart and Hansche[11] at somewhat lower Mach numbers.

In order that measured drag coefficients of steel pellets might be compared with those obtained from projectiles other than spherical or cylindrical, Eq. (7) is written in the form

$$a = -K_D \rho \left(\frac{4}{\pi}\right)\left(\frac{A}{m}\right) v^2, \tag{8}$$

where m is the mass of the projectile and A is its presentation area.

Comparison of Eq. (2) and Eq. (8) yields the following relationship between α and K_D

$$K_D = \alpha(\pi/4)(m/A)(1/\rho). \tag{9}$$

The average presentation area of a number of recovered steel pellets have been measured and found to average 0.680 cm². If this value and the appropriate value of α listed in Table II are substituted in Eq. (9), a value of K_D equal to 0.328 is obtained. Charters and Thomas report a value of K_D equal to about 0.38 for a sphere of $\frac{9}{16}$-in. diameter at the velocity of 1.0 m/msec. Rinehart and Hansche's value of K_D for cubes is about 50 percent higher than the K_D for spheres. Data for a simple projectile with a conical head are given by Thomas,[12] where measured values of K_D range as low as 0.12 for Mach number 2.550.

The Charters-Thomas data for spheres and the values of K_D obtained from the present ultra-speed pellet studies are plotted in Fig. 6 for purposes of comparison. Charters and Thomas experimented with spheres of two different diameters, $\frac{9}{16}$ in. and $\frac{9}{32}$ in. The volume of the pellet is equivalent to a sphere 0.7860-cm diameter and 0.4852-cm² cross section. From Table III it can be seen that both the volume of the steel pellet and its average presentation area lie between those of the two spheres. For a given Mach number, Charters and Thomas report no essential difference between the values of K_D for spheres of $\frac{9}{16}$-in. diameter and $\frac{9}{32}$-in. diameter.

BURNING PELLETS

It has been noted, see Eq. (1), that all ballistic data pertaining to burning pellets can be represented within

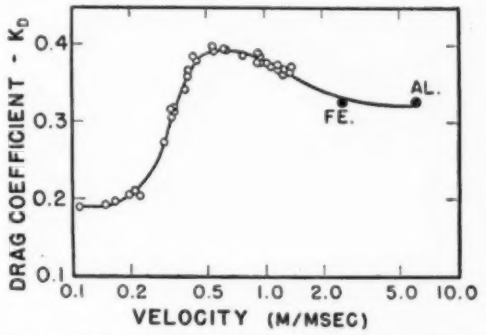


FIG. 6. Values of K_D for steel and aluminum pellets compared with the Charters-Thomas data on spheres.

experimental accuracy by the relation

$$v = v_0 e^{-\alpha s}. \tag{10}$$

This relation seems to hold whether or not the pellet breaks up. The value of v_0 for a given material is constant, while the value of α seems to be a function of the number of pieces into which the pellet breaks up. For pellets that break up excessively, the measured value of α is relatively large. Accordingly, it is believed that the smallest measured value of α for a given material most nearly represents the performance of an ultra-speed pellet that remains intact. In the compilation of Table II the value listed for α represents an envelope of all measured values. Owing to present paucity of experimental data, no probable errors have been calculated for the values listed in Table II. An idea of the experimental scatter can be inferred from Fig. 4 where the experimental data for four magnesium-lithium pellets have been plotted. A straight line has been fitted by the method of least squares to the data of each pellet. It might be inferred from Fig. 4 that the data indicate a sinusoidal departure from this best fitting straight line. No systematic deviation, however, is believed to exist even at high Mach numbers. The values of v_0 and α listed in Table II have been plotted as functions of pellet density in Fig. 7. The point U in Fig. 7 represents the detonation velocity and the point c represents 0.8 of the detonation velocity of the high explosive. According to Penney,[13] the latter value is a resaonable estimate of the sound velocity behind the detonation front. It may be that the sound velocity of the high explosive might represent a theoretical upper limit to the attainable pellet velocity. The value of α

TABLE III. Comparison of pellet volume and presentation area with those of spheres of comparable mass.

Projectile	Volume (cc)	Average presentation area (cm²)
$\frac{9}{16}$-in. diameter sphere (*C* & *T*)	1.527	1.603
Steel pellet (NOTS)	0.254	0.680
$\frac{9}{32}$-in. diameter sphere (*C* & *T*)	0.191	0.401

[10] A. C. Charters and R. N. Thomas, J. Aeronaut. Sci. **12**, 468 (1945).
[11] J. S. Rinehart and G. E. Hansche, Pop. Astron. **58**, 308 (1950).
[12] R. N. Thomas, "The ballistic contribution to meteorics," Harvard Reprint No. 34 Series II (1950).
[13] W. G. Penney, Proc. Roy. Soc. (London) **204**, 3 (1950).

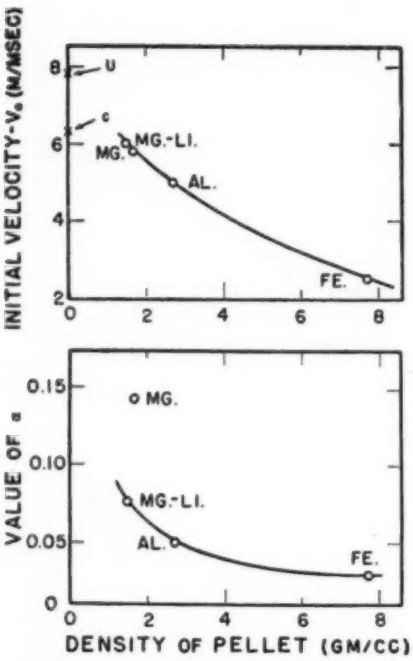


FIG. 7. Variation of α and v_0 in the equation, $v=v_0e^{-\alpha x}$, as a function of pellet density. The point U represents the detonation velocity of the high explosive; the point c is a measure of the sound velocity of the detonation products.

plotted in Fig. 7 for magnesium pellets seems to be anomalous; this anomaly probably arises from the fact that magnesium does not flow as well as the other materials and hence tends to break up.

The experimental data can be fitted reasonably well with a straight line on semi-log paper and Eq. (10) can be written in the form

$$v=v_0e^{-K_D\rho(4/\pi)(A/m)x}. \tag{11}$$

Therefore, it can be inferred that the quantity K_DAm^{-1} must be constant, at least to a first approximation. Reference to Fig. 6 indicates that a deformed steel pellet has approximately the same value of K_D as a sphere of the same mass; therefore, K_D must be a slowly varying function of shape. It follows that the quantity Am^{-1} must remain essentially constant during the burning process. Such would be the case for a pellet that has been flattened into a disk by the action of the explosive and subsequently suffers erosion along its edges. This possibility is not incompatible with the shape of recovered pellets, nor with the presentation areas as measured with yaw cards.

The calculated value of K_D for the steel pellets can be used to determine the initial presentation area A_0 of the burning pellets, provided it is assumed that K_D is a slowly varying function of velocity for this region of the ultra-speed regime. The initial presentation area is given by the equation

$$A_0=\frac{m_0\alpha\pi}{4\rho K_D}, \tag{12}$$

where m_0 represents the initial mass of the pellet. Values calculated from Eq. (12) are tabulated in Table IV.

AERODYNAMIC CONSIDERATIONS

A spherical projectile moving through air at supersonic velocities is preceded by a strong head wave. In general, any object moving through gas at supersonic speed generates a shock wave. The shock wave can be regarded as a surface of discontinuity separating those parts of the gas that have been disturbed by the body from those that have not. Sound waves in subsonic flow travel faster than the object, and signal to the entire medium the presence of the object with sufficient warning to allow the thermodynamic state of the medium to adjust itself. In supersonic flow, however, the thermodynamic state of the medium is changed abruptly across the shock. For a sufficiently small disturbance, this change in thermodynamic state can be effected in a few mean free paths, and neglect of the shock wave thickness is justified. In this case, the state of the medium behind the shock wave can be determined from the state of the undisturbed medium, from a knowledge of the shock wave velocity, and from the conservation laws of mass, momentum, and energy. For a blunt missile the shock wave velocity at the tip of the missile is identical with that of the missile. For a missile velocity of about 6 m/msec, the temperature immediately behind the shock wave is, on the basis of classical aerodynamical calculations, about 24,000°K. The calculations based on classical aerodynamics cannot represent physical reality because the simple theory neglects, among other effects, the variation of γ across the shock wave. This variation can be introduced into the calculations if the thermodynamic states are known at all points within the shock wave. These intermediate states are not states of thermodynamic equilibrium, at least for very strong shock waves. A more realistic calculation must involve the rate of activation of the various degrees of freedom of the gas.

For nonburning pellets it has been determined that the pellets are nonluminous at velocities below 2.4 m/msec. This value corresponds roughly to the velocity at which a normal shock wave ceases to be luminous. This fact was inferred from data obtained from photographs of the atomic explosion of 1945.[14]

TABLE IV. Average initial presentation areas of pellets.

Pellet material	A_0 (cm²)
Steel	0.68
2S—O aluminum	0.68
Magnesium	1.26[a]
Mg—Li alloy	0.58

[a] It is suggestive that this presentation area is equal to the original area of the ½-in. disk; i.e., 1.267 cm².

[14] G. I. Taylor, Proc. Phys. Soc. (London) **201**, 178 (1950).

CONCLUSIONS

The data presented in this report are of a preliminary nature and may be refined as a result of further experimental work. At present, drag determinations on the flight of ultra-speed pellets do not indicate any significant deviation of the drag coefficient from accepted values extrapolated to large Mach numbers. Further study of ultra-speed pellets might contribute to the problem of physical states not in thermodynamic equilibrium.

It is a pleasure to acknowledge the assistance of the members of this group who contributed to the work reported herein, including Mr. Clyde L. McCrary, Mr. J. Craig McLaughlin, Mr. R. D. Smith, and Mr. Joseph Wojecki. In addition, appreciation is felt for the assistance of Mr. G. C. Throner and other members of the Ordnance Section, U.S. Naval Ordnance Test Station.

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 1 JANUARY, 1952

The Probability Distribution of the Phase of the Resultant Vector Sum of a Constant Vector Plus a Rayleigh Distributed Vector

KENNETH A. NORTON, EDNA L. SHULTZ, AND HELEN YARBROUGH
National Bureau of Standards, Boulder, Colorado

(Received August 22, 1951)

Formulas, tables, and graphs are given of the cumulative probability distribution of a function frequently occurring in the theory and practice of radio wave propagation as well as in the study of the influence of noise in phase modulation systems.

SUPPOSE the x and y coordinates of a point on a plane to be independent normally distributed random variables with the same variance and zero means. Let $E_s(t)$ and $\theta(t)$ denote the polar coordinates of the point. (See Fig. 1.) It is well known that the distance, $E_s(t)$, from the origin is then Rayleigh distributed while all values of its angular position, $\theta(t)$, between 0 and 2π are equally likely. Now let the mean value of the x coordinate have a finite value, E_0, and consider the polar coordinates, $E(t)$, $\phi(t)$, of the point relative to the displaced origin.

The probability distribution of the distance, $E(t)$, of the point from the displaced origin is already available in the literature in convenient form for radio applications.[1,2] The present paper deals with the probability distribution of the angular position, $\phi(t)$, of the point relative to the displaced origin. Solutions of this problem have been obtained by statisticians[3–8] in connection with other applications. The present derivation is carried out from a somewhat different point of view, and the accompanying tables and graphs are presented in forms convenient for the radio applications contemplated.

The need for a solution of this two-dimensional random walk problem arises in many phases of the theory and practice of radio wave propagation as well as in the study of various modulation systems, e.g., noise in phase modulation systems. As a particular example, consider a radio navigation system which depends on the measurement of the phase of the ground wave for a position determination. At moderate distances from the transmitting antenna, waves scattered from the ionosphere perturb the phase of the ground wave in a random fashion, and it is desirable to determine the percentage of time that this perturbation in phase may be expected to exceed various specified values. It is known from measurements that there are occasions when all values of the phase of the scattered ionospheric wave are equally likely while its amplitude is Rayleigh distributed.[9] Thus, if we denote by E_r, the root-mean-square value of the amplitude of this scattered wave, $E_s(t)$, then the probability that its instantaneous amplitude,*[10] $E_s(t)$,† will exceed rE_0 is

[1] S. O. Rice, Bell System Tech. J. **23**, 282–332 (1944); **24**, 46–156 (1945).
[2] Kenneth A. Norton, "Propagation in the FM broadcast band," *Advances in Electronics* (Academic Press, Inc., New York, 1948), Vol. I, pp. 406–408.
[3] W. F. Sheppard, Trans. Cambridge Phil. Soc. **19**, 24–69 (1900).
[4] C. Nicholson, Biometrika **33**, 59–72 (April 1943).
[5] H. H. Germond, "Miscellaneous Probability Tables," AMP Note No. 14, Applied Mathematics Panel, NDRC, OSRD (July 1944).
[6] R. C. Geary, J. Roy. Stat. Soc. **93**, 442–446 (1930).
[7] E. C. Fieller, Biometrika **24**, 428–440 (November 1932).
[8] C. Nicholson, Biometrika **32**, 16–28 (January 1941).
[9] Ionospheric Radio Propagation, National Bureau of Standards Circular 462, p. 108, U. S. Government Printing Office (June 1948).
* It is important to distinguish the instantaneous amplitude, $E_s(t)$, of the envelope of the scattered wave from its instantaneous voltage, $v_s(t)$, to which it is related by the following equation:

$$v_s(t) = E_s(t)\sin\omega t.$$

It has been shown (see reference 10) that $v_s(t)$ is normally distributed about a mean of zero with a variance equal to $E_r^2/2$; in other words, the mean square amplitude, E_r^2, is just twice the value of the mean square voltage.
[10] Kenneth A. Norton, Proc. Inst. Radio Engrs. **30**, 425–429 (1942).
† $E_s(t)$ is the notation adopted here for a random variable, the variations of which take place with respect to time in applications to ionospheric wave propagation, but which may take place with respect to other variables in other applications.

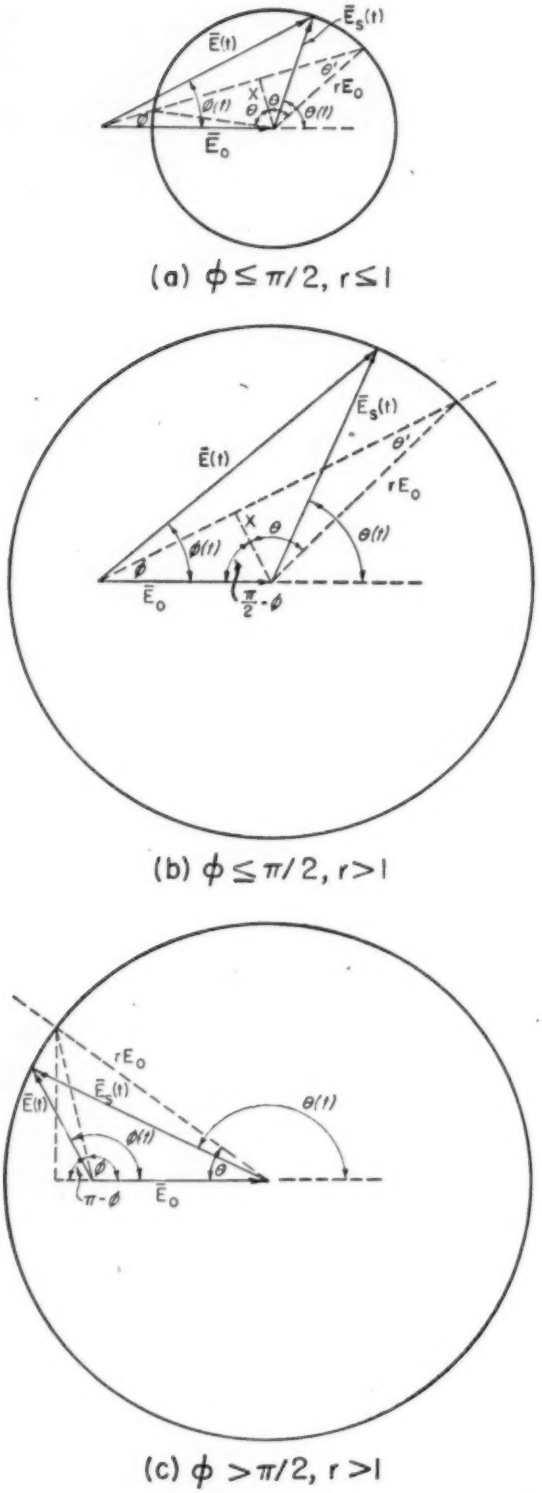


FIG. 1. The phase, ϕ, of the resultant vector sum of a constant vector, $\bar{E}_0$, plus a Rayleigh distributed vector, $\bar{E}_s$.

given by

$$p[E_s(t)>rE_0]=\exp(-r^2E_0^2/E_r^2)\equiv\exp(-r^2/k),$$

where $k\equiv E_r^2/E_0^2$. Let $\phi(t)$ denote the phase of the resultant vector $\bar{E}(t)$ relative to that of the ground wave, $\bar{E}_0$, as shown on Fig. 1. Our problem is to determine the cumulative probability distribution of $\phi(t)$, i.e., the probability that $\phi(t)$ will exceed ϕ. In the first place, it should be noted that positive and negative values of $\phi(t)$ are equally likely, so we may confine our attention to the values of $\phi(t)$ between 0 and π. Since all values of the phase, $\theta(t)$, of the scattered Rayleigh-distributed wave are taken to be equally likely, the terminal point of $\bar{E}_s(t)$ as well as the terminal point of the resultant vector, $\bar{E}(t)=\bar{E}_0+\bar{E}_s(t)$ will be with equal likelihood at any point on the circle of radius, E_s, when the amplitude of the Rayleigh distributed vector is equal to E_s. From the geometry of Fig. 1(a) or 1(b)

$$X=E_0\sin\phi=E_s\cos\theta=rE_0\cos\theta \tag{2}$$

$$\therefore\quad r=\sin\phi/\cos\theta. \tag{3}$$

For the particular value of r shown on Fig. 1(a), we see that the conditional probability, $q[|\phi(t)|>\phi|r]$, that $\phi(t)$ will exceed ϕ is simply equal to $2\theta/\pi$, but for the particular value of r shown on Fig. 1(b), i.e., $r>1$ and $\theta>(\pi/2)-\phi$, $q[|\phi(t)|>\phi|r]=(\theta+(\pi/2)-\phi)/\pi$. Reference to Fig. 1(a) shows that $q=0$ when $r=\sin\phi$ and $q=1$ when $r=\infty$. Since $E_s(t)$ and $\theta(t)$ are independent random variables, q and dp will also be independent and the probability, $P[|\phi(t)|>\phi]$, that $\phi(t)$ will exceed ϕ may be obtained for the case $\phi<\pi/2$ by summing the products of these independent probabilities from $r=\sin\phi$ to $r=\infty$:

$$P[|\phi(t)|>\phi]=-\int_{r=\sin\phi}^{r=\infty}q(dp/dr)dr=\int_{r=\sin\phi}^{r=\infty}pdq$$

$$=(2/\pi)\int_0^{(\pi/2)-\phi}pd\theta+(1/\pi)\int_{(\pi/2)-\phi}^{\pi/2}pd\theta$$

$$=(2/\pi)\int_0^{\pi/2}pd\theta-(1/\pi)\int_{(\pi/2)-\phi}^{\pi/2}pd\theta. \tag{4}$$

Upon introducing (3) into (1) and the result into (4) and after replacing θ by $(\pi/2)-\theta'$ in the second integral, we obtain

$$P[|\phi(t)|>\phi]=(2/\pi)\int_0^{(\pi/2)}\exp(-\sin^2\phi/k\cos^2\theta)d\theta$$

$$-(1/\pi)\int_0^{\phi}\exp(-\sin^2\phi/k\sin^2\theta')d\theta', \tag{5}$$

where $\phi<\pi/2$. From Fig. 1(c) we may write

$$\tan(\pi-\phi)=-\tan\phi=r\sin\theta/(r\cos\theta-1) \tag{6}$$

$$\therefore\quad r=\sin\phi/\sin(\theta+\phi). \tag{7}$$

The probability that $\phi(t)$ will exceed ϕ is

$$P[|\phi(t)|>\phi]=(1/\pi)\int_0^{\pi-\phi}pd\theta$$

$$=(1/\pi)\int_0^{\pi-\phi}\exp[-\sin^2\phi/k\sin^2(\theta+\phi)]d\theta.$$

TABLE I. The coefficients in Eq. (16).

k	A	B	C
0.1	0.000001	0.000000	0.000000
0.2	0.000170	0.000074	0.000049
0.5	0.006775	0.002663	0.001633
1.0	0.028353	0.010066	0.005763
2.0	0.066476	0.021098	0.011315
5.0	0.127618	0.034927	0.017450
10.0	0.171208	0.042296	0.020333
20.0	0.207937	0.046999	0.022004
50.0	0.244867	0.050369	0.023098
100.0	0.265069	0.051640	0.023481

Let $\theta' = \pi - \theta - \phi$

$$P[|\phi(t)|>\phi] = (1/\pi)\int_0^{\pi-\phi} \exp(-\sin^2\phi/k\,\sin^2\theta')d\theta', \quad (8)$$

where $\phi > \pi/2$. It may be shown that the first integral in (5) is simply the error function of $(\sin\phi/(k)^{\frac{1}{2}})$ and thus (5) may be expressed

$$P[|\phi(t)|>\phi] = erfc(\sin\phi/(k^{\frac{1}{2}})) - (1/\pi)\int_0^{\phi} \exp(-\sin^2\phi/k\,\sin^2\theta')d\theta', \quad (9)$$

where $\phi < \pi/2$.

$$erfc(x) \equiv (2/\pi^{\frac{1}{2}})\int_x^{\infty} \exp(-u^2)du = 1-(2/\pi^{\frac{1}{2}})\int_0^{x} \exp(-u^2)du. \quad (10)$$

It is interesting to note that the integral in (8) is of the same form as the second integral in (9) and that these integrals are identical when $\phi = \pi/2$. After the first draft of this paper was completed, it was found that the second integral in (9), which will be designated $Q(k, \phi)$, may be evaluated as follows:

$$Q(k,\phi) = \phi/\pi - \tfrac{1}{2}erfc(\sin\phi/k^{\frac{1}{2}}) + 2V(h,q), \quad (11)$$

where $h = (2/k)^{\frac{1}{2}}\sin\phi$ and $q = (2/k)^{\frac{1}{2}}\cos\phi$ and $V(h, q)$ is a function tabulated by Nicholson[4] to 6 decimals for $h, q = 0(.1)3$, and by Germond[5] to 5 decimals for $h = 0(.01)4$ and $q = 0(.1)1$. In the present paper graphical methods were used for the evaluation of this integral in those cases where k is small. At the other extreme, for very large values of k, the exponential factor approaches unity and we obtain from either (8) or (9)

$$\operatorname*{Limit}_{k\to\infty} P[|\phi(t)|>\phi] = 1-(\phi/\pi). \quad (12)$$

It should be noted that an infinitely large value of k corresponds simply to another Rayleigh distribution for which all values of the phase are equally likely and (12) is simply a statement of this conclusion.

For the particular case $\phi = (\pi/2)$ it is easy to see from (8) and (9) that

$$P[|\phi(t)|>\pi/2] = \tfrac{1}{2}erfc[1/(k)^{\frac{1}{2}}]. \quad (13)$$

When ϕ or $(\pi-\phi)$ are small, the following development of the second integral in (9) is useful:

$$Q(k,\phi) = (1/\pi)\int_0^{\phi} \exp(-\sin^2\phi/k\,\sin^2\theta')d\theta' \quad (14)$$

$$\text{set } u^2 = \sin^2\phi/k\,\sin^2\theta'$$

$$Q(k,\phi) = \frac{\sin\phi}{\pi(k^{\frac{1}{2}})}\int_{[1/(k)^{\frac{1}{2}}]}^{\infty} \frac{\exp(-u^2)du}{u^2[1-(\sin^2\phi/u^2k)]^{\frac{1}{2}}}$$

$$= \frac{\sin\phi}{\pi(k^{\frac{1}{2}})}\int_{[1/(k)^{\frac{1}{2}}]}^{\infty} \exp(-u^2)\left[\frac{1}{u^2}+\frac{\sin^2\phi}{2ku^4} + \frac{3\sin^4\phi}{8k^2u^6}+\cdots\right]du, \quad (15)$$

$$Q(k,\phi) = \sin\phi A + \sin^3\phi B + \sin^5\phi C\cdots, \quad (16)$$

TABLE II. The cumulative probability $P[|\phi(t)|>\phi]$.

ϕ	$k=0.01$	0.02	0.05	0.1	0.2	0.5	1	2	5	10	20	50	100	∞
0.1	0.98031	0.98608	0.99119	0.99377	0.99560	0.99720	0.99798	0.99849	0.99890	0.99908	0.99920	0.99929	0.99934	0.99944
0.2	0.96063	0.97215	0.98239	0.98754	0.99119	0.99440	0.99596	0.99698	0.99779	0.99816	0.99839	0.99859	0.99868	0.99889
0.5	0.90178	0.93046	0.95599	0.96887	0.97798	0.98602	0.98991	0.99246	0.99448	0.99539	0.99598	0.99647	0.99670	0.99722
1.0	0.80505	0.86145	0.91211	0.93779	0.95598	0.97204	0.97981	0.98492	0.98897	0.99078	0.99197	0.99294	0.99340	0.99444
2.0	0.62162	0.72709	0.82531	0.87597	0.91212	0.94412	0.95965	0.96984	0.97794	0.98157	0.98394	0.98588	0.98681	0.98889
3.0	0.45921	0.60072	0.74064	0.81494	0.86854	0.91628	0.93951	0.95478	0.96691	0.97236	0.97591	0.97883	0.98021	0.98333
4.0	0.32388	0.48545	0.65908	0.75507	0.82540	0.88857	0.91944	0.93974	0.95590	0.96316	0.96788	0.97177	0.97362	0.97778
5.0	0.21774	0.38345	0.58148	0.69670	0.78283	0.86103	0.89943	0.92474	0.94490	0.95396	0.95986	0.96472	0.96703	0.97222
10.0	0.01406	0.08248	0.27210	0.43741	0.58289	0.72718	0.80103	0.85049	0.89020	0.90814	0.91985	0.92951	0.93410	0.94444
15.0	0.00025	0.00965	0.10165	0.24708	0.41305	0.60291	0.70683	0.77819	0.83632	0.86278	0.88011	0.89444	0.90127	0.91667
20.0	0.00000	0.00063	0.03053	0.12613	0.27939	0.49152	0.61848	0.70870	0.78362	0.81809	0.84077	0.85959	0.86859	0.88889
30.0	0.00000	0.00000	0.00157	0.02535	0.11375	0.31353	0.46388	0.58078	0.68301	0.73142	0.76371	0.79076	0.80378	0.83333
40.0	0.00000	0.00000	0.00005	0.00404	0.04195	0.19327	0.34155	0.47032	0.59043	0.64945	0.68953	0.72351	0.74000	0.77778
50.0	0.00000	0.00000	0.00000	0.00061	0.01523	0.11836	0.24980	0.37827	0.50706	0.57306	0.61882	0.65823	0.67753	0.72222
60.0	0.00000	0.00000	0.00000	0.00011	0.00592	0.07384	0.18342	0.30353	0.43322	0.50271	0.55201	0.59518	0.61656	0.66667
70.0	0.00000	0.00000	0.00000	0.00003	0.00262	0.04785	0.13615	0.24381	0.36857	0.43849	0.48879	0.53455	0.55723	0.61111
80.0	0.00000	0.00000	0.00000	0.00001	0.00134	0.03227	0.10235	0.19642	0.31239	0.38019	0.43064	0.46643	0.49962	0.55556
90.0	0.00000	0.00000	0.00000	0.00000	0.00078	0.02275	0.07865	0.15866	0.26354	0.32736	0.37591	0.42074	0.44377	0.50000
100.0	0	0	0	0	0.00051	0.01661	0.06135	0.12830	0.22100	0.27944	0.32484	0.37742	0.38962	0.44444
110.0	0	0	0	0	0.00034	0.01234	0.04772	0.10357	0.18373	0.23582	0.27756	0.31638	0.33705	0.38889
120.0	0	0	0	0	0.00026	0.00943	0.03725	0.08294	0.15067	0.19583	0.23218	0.26731	0.28597	0.33333
130.0	0	0	0	0	0.00019	0.00714	0.02885	0.06538	0.12098	0.15885	0.18977	0.22001	0.23620	0.27778
140.0	0	0	0	0	0.00014	0.00532	0.02178	0.05005	0.09392	0.12431	0.14940	0.17420	0.18757	0.22222
150.0	0	0	0	0	0.00010	0.00378	0.01562	0.03630	0.06882	0.09164	0.11066	0.12958	0.13985	0.16667
160.0	0	0	0	0	0.00006	0.00243	0.01013	0.02363	0.04513	0.06034	0.07310	0.08587	0.09284	0.11111
170.0	0	0	0	0	0.00003	0.00119	0.00498	0.01166	0.02235	0.02996	0.03636	0.04279	0.04630	0.05556
180.0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.00000

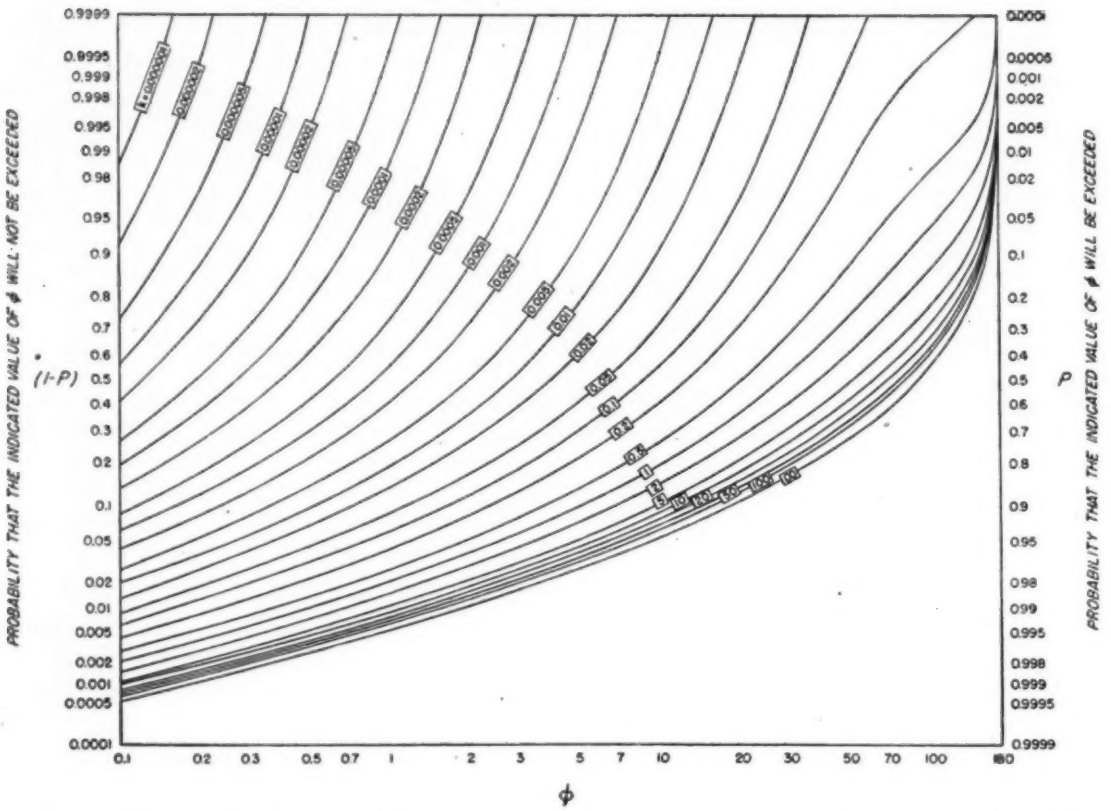

PROBABILITY THAT THE INDICATED VALUE OF φ WILL NOT BE EXCEEDED
(1-P)
PROBABILITY THAT THE INDICATED VALUE OF φ WILL BE EXCEEDED
P
φ

FIG. 2. The cumulative distribution of the phase, ϕ, of the resultant vector sum of a constant vector plus a Rayleigh distributed vector ($k=E_r^2/E_0^2$; E_r^2=mean square amplitude of Rayleigh distributed vector; E_0=amplitude of constant vector).

where

$$A=\frac{1}{\pi(k^{\frac{1}{2}})}\int_{[1/(k)^{\frac{1}{2}}]}^{\infty}[\exp(-u^2)/u^2]du$$

$$=(1/\pi)\{\exp(-1/k)-(\pi/k)^{\frac{1}{2}}erfc[1/(k)^{\frac{1}{2}}]\} \quad (17)$$

$$B=\frac{1}{2\pi k(k)^{\frac{1}{2}}}\int_{[1/(k)^{\frac{1}{2}}]}^{\infty}\frac{\exp(-u^2)}{u^4}du=\frac{1}{6\pi}\left\{\left(1-\frac{2}{k}\right)\right.$$

$$\left.\times\exp(-1/k)+2(\pi/k^3)^{\frac{1}{2}}erfc[1/(k)^{\frac{1}{2}}]\right\} \quad (18)$$

$$C=\frac{3}{8\pi k^2(k^{\frac{1}{2}})}\int_{[1/(k)^{\frac{1}{2}}]}^{\infty}\frac{\exp(-u^2)}{u^6}du=\frac{3}{40\pi}\left\{\left(1-\frac{2}{3k}+\frac{4}{3k^2}\right)\right.$$

$$\left.\times\exp(-1/k)-\frac{4}{3}(\pi/k^5)^{\frac{1}{2}}erfc[1/(k)^{\frac{1}{2}}]\right\}. \quad (19)$$

The coefficients A, B, and C are given in Table I for several values of k. By using Table I in (16), Q may be calculated to an accuracy of 5 decimal places for angles $\phi<18°$ for any value of k.

Finally by setting $u^2=(v+1)/k$ in (14) we obtain

$$Q(k,\phi)=\frac{\tan\phi\exp(-1/k)}{2\pi}\times\int_0^{\infty}\frac{\exp(-v/k)dv}{(1+v)[1+(v/\cos^2\phi)]^{\frac{1}{2}}} \quad (20)$$

$$Q(k,\phi)=\frac{\tan\phi\exp(-1/k)}{2\pi}\int_0^{\infty}\exp(-v/k)[1-v+v^2-\cdots]\left[1-\frac{v}{2\cos^2\phi}+\frac{1\cdot3v^2}{2\cdot4\cos^4\phi}-\cdots\right]dv \quad (21)$$

$$Q(k,\phi)=\frac{k\tan\phi\exp(-1/k)}{2\pi}\left[1-k\left\{1+\frac{1}{2\cos^2\phi}\right\}\right.$$

$$\left.+k^2\left\{2+\frac{1}{\cos^2\phi}+\frac{3}{4\cos^4\phi}\right\}-\cdots\right]. \quad (22)$$

When k is sufficiently small, the above may be used for fairly large values of ϕ, i.e., such that $[(\pi/2)-\phi]\gg k$; for comparison with the above, the corresponding expansion for small k when $\phi=\pi/2$ may be obtained directly from the well-known asymptotic expansion of

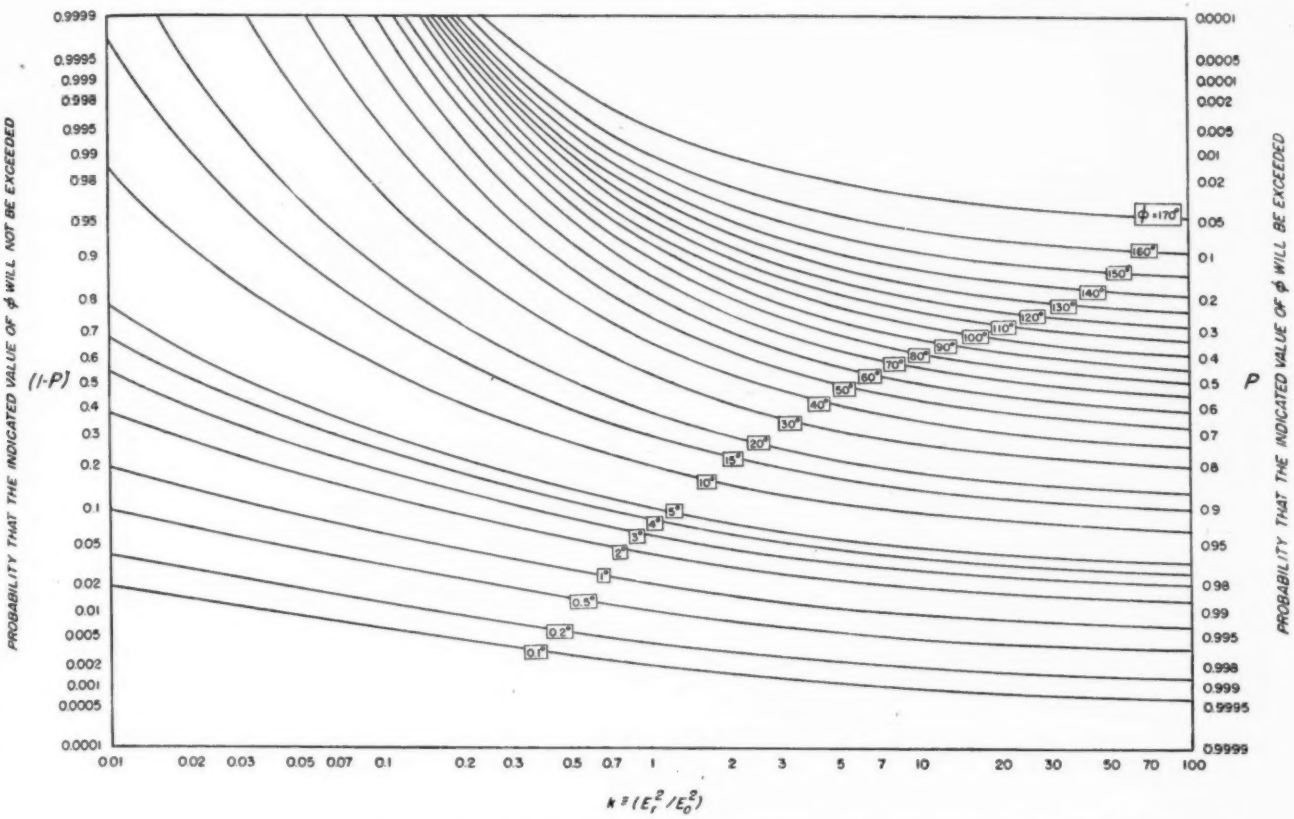


FIG. 3. The cumulative distribution of the phase, ϕ, of the resultant vector sum of a constant vector plus a Rayleigh distributed vector ($k=E_r^2/E_0^2$; E_r^2=mean square amplitude of Rayleigh distributed vector; E_0=amplitude of constant vector).

the error function

$$P[|\phi(t)|>\pi/2]=[(\pi k)^{\frac{1}{2}}\exp(-1/k)/2\pi]\times[1-(k/2)+(1\cdot 3k^2/4)-(1\cdot 3\cdot 5k^3/8)+\cdots]. \quad (23)$$

P is tabulated as a function of ϕ and k in Table II. In determining the values in this table, (14) was integrated numerically in those cases where none of the above formulas was sufficiently accurate to determine P to five decimal places. $(1-P)$ and P are also shown graphically on Figs. 2 and 3.

ACKNOWLEDGMENT

Acknowledgment is made of the helpful suggestions of J. J. Freeman relative to the method of presentation.

Steady-State One-Dimensional Flow in Rocket Motors

E. W. PRICE
Michelson Laboratory, U. S. Naval Ordnance Test Station, Inyokern, China Lake, California
(Received August 22, 1951)

Solutions are presented for the problem of gas flow in a solid propellant rocket motor using combined numerical and graphical methods. The character of the problem, sources of information, and method of solution are described, and solutions for the case of JPN ballistite propellant are presented. These solutions give the equilibrium distributions of pressure, density, velocity, and temperature of the gas stream and burning rate of the solid propellant as functions of two characteristic parameters β and J. These parameters are descriptive of the charge geometry and flow channel geometry, respectively. The solutions cover the entire range of nozzle, charge-size combinations, including the case of no nozzle.

INTRODUCTION

DURING the development of a science of interior ballistics for rocket motors, the engineer has been continually plagued by the lack of a clear statement and solution of the problem of gas flow in the motor. This deficiency in the theoretical developments is due not to any essential difficulty in the flow problem, but to the need to consider a large number and variety of physical and thermodynamic variables. This inconvenience is particularly significant in the case of rockets using solid propellant charges which burn on the longitudinal surfaces, where flow adjacent to the burning surface must be considered.

This paper presents a solution to the flow problem in solid propellant rocket motors, using a combined numerical and graphical method. The solution involves a comparatively small number of assumptions, and is particularly adaptable to utilization of empirically determined properties of the solid propellant and propellant gas.

The results of the type presented are essential to accurate prediction of motor thrust, to rational stress calculation for motor tubes, to prediction of heat transfer to metal components, to prediction of operating and design limitations and to optimization of rocket motor design.

STATEMENT OF THE PROBLEM

The arrangement of the components in a typical solid propellant rocket motor is indicated by the diagram in Fig. 1. During operation, the solid charge burns progressively inward from the surface, and the gas evolved flows along the charge and out through the nozzle. The rate at which a particular propellant burns away at the surface has been shown to depend primarily upon the initial temperature T_p of the solid propellant, the static pressure p, and the velocity u of the adjacent stream.[1] The mass flux m in the channel at a given cross section is proportional to the integral of the burning rate over the charge surface upstream of the point in the channel. The pressure, velocity, density ρ, and temperature T of the gas at each point in the flow channel are given by the laws of conservation of energy, mass and momentum, and the equation of state for the gas.

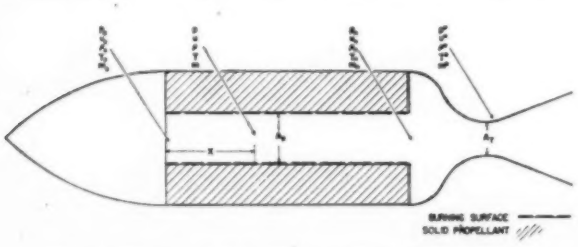


FIG. 1. Diagram of a solid propellant rocket motor.

It will be shown that, by choosing a pressure p_0 at the upstream end of the channel as a boundary condition, the equations of flow and of burning permit calculation of a flow field in the channel adjacent to the charge corresponding to an equilibrium between the interactions of the burning rate and the flow field. Variation in the choice of the pressure p_0 corresponds to generation of a family of solutions to the flow field in the motor, variation in length of the charge determines what portion of the flow field is actually obtained, and variation in the nozzle size determines which solution (p_0) is actually obtained.

INITIAL ASSUMPTIONS: ANALYTICAL STATEMENT OF PROBLEM

In order to reduce the flow problem to tractable character the flow process will be approximated by a one-dimensional, steady-state, adiabatic nonviscous flow, so that the variables of state and the burning rate are functions of the distance x from the upstream end of the propellant charge. Consideration will be given only to channels whose cross-sectional shapes are independent of x. These approximations are discussed by Wimpress[2] and do not appear to restrict the applicability of the solutions unduly.

Using the assumptions made above, the equations of state, continuity, momentum, and energy have the

[1] R. N. Wimpress, *Internal Ballistics of Solid-Fuel Rockets* (McGraw-Hill Book Company, Inc., New York, 1950), pp. 16 17, 22.
[2] See reference 1, pp. 29, 52–57.

forms

$$p=\rho(RT/\mu), \tag{1}$$

$$m=\rho u A_p, \tag{2}$$

$$\rho u^2=p_0-p, \tag{3}$$

$$h+(\tfrac{1}{2})u^2=h_0, \tag{4}$$

where T is the temperature of the gas in the stream, μ is the molecular weight, R is the universal gas constant, A_p is the cross-sectional area of the flow channel adjacent to the propellant charge, h is the specific enthalpy of the propellant gases. The specific stagnation enthalpy h_0 is practically constant for any one propellant, and equal to the specific enthalpy at the upstream end of the flow channel.

The molecular weight μ and specific enthalpy h can be determined as functions of propellant composition, the temperature T and pressure p, by solution of the equations for chemical equilibrium. A method of solution of these equations is described by McEwan and Skolnik,[3] and additional references to the problem are noted in their paper. The results of solution of the equations of chemical equilibrium are usually presented in graphical form as in Figs. 2.2 and 2.3 of reference 1. Figure 2.2 shows the dependence of molecular weight on temperature for curves of constant pressure, and Fig. 2.3 shows the dependence of temperature on specific entropy for curves of constant enthalpy and curves of constant pressure. The burning rate r is determined by direct observation in closed bombs or by interrupted or complete reaction of measured charges in rocket motors. The data obtained are highly reproducible for most usable propellants, and in many cases are found to fit the relation

$$r=C(T_p)f(p)g(u)=Cp^n(1+ku), \tag{5}$$

where C, n and k are experimentally determined parameters for any one propellant, C being dependent upon the initial temperature T_p of the solid propellant charge. Regardless of the dependency of the burning rate, the mass flux m in a channel of area A_p is given by the relation

$$(m/A_p)=(\rho_p q/A_p)\int_0^x r\,dx, \tag{6}$$

where ρ_p is the density of the solid propellant and q is the perimeter of the cross section of the burning surface.

In the event that Eq. (5) or some other dependency (empirical or functional) of burning rate on variables of state is available, Eqs. (1)–(6) and Figs. 2.2 and 2.3 of reference 1 define a set of eight relations among nine variables h, m, p, r, T, u, ρ, μ, and x, where the quantities A_p, ρ_p, q, C, n, k, and h_0 are determined by the choice of propellant, of charge shape, and of propellant temperature. Thus it should be possible to solve for the first eight variables in terms of x.

[3] W. S. McEwan and Sol Skolnik, Rev. Sci. Instr. **22**, 125 (1951).

For a rocket motor with a nozzle constriction downstream of the propellant charge, the value of p_0 obtained in practice is determined by the requirement that the mass flux through the nozzle must equal the mass flux from the burning surface of the charge for steady-state operation. Since the discharge rate of a smooth-contoured nozzle with sonic flow at the throat can be expressed in terms of the pressure (stagnation) at the nozzle entry and the area of the nozzle throat,[4,5] the equilibrium between the above mass rates is accomplished by matching the mass burning rate and downstream pressure for the problem of flow in the charge channel with the pressure-discharge characteristic for the nozzle used. This process uniquely determines the flow field in the charge channel. The flow field in the nozzle can then be determined by the one-dimensional theories familiar in wind-tunnel and other flow channel designs.[4–6]

NUMERICAL SOLUTION

In view of the numerical form of the thermodynamic and burning rate data on propellants, and the integral form of Eq. (6), a numerical method of solution of the flow problem is advantageous. To this end, Eq. (6) may be written in the form

$$(m/A_p)=(\rho_p q/A_p)\int_0^x r\,dx=\sum \overline{f(p)}\,\overline{g(u)}\,\Delta\beta, \tag{7}$$

where $\bar{f}$ and $\bar{g}$ are average values of f and g for the interval $\Delta\beta$, and

$$\beta=(\rho_p q C/A_p)x. \tag{8}$$

The factors entering into the parameter β all affect the solution to Eq. (7) in the same way, so that the solutions can be materially condensed without loss in content by presenting them in terms of β. In the sense of a position x along the flow channel, β will be used interchangeably as a position on a given charge (hence giving space distributions of related variables) and as the over-all length of a charge (hence giving dependence of solutions on length of a charge).

Using the initial conditions of $u=0$, $h=h_0$, $p=p_0$ at $x=0$, the summation in Eq. (7) was performed by a sequence of calculations successively using Fig. 2.3 of reference 1 for T, Fig. 2.2 of reference 1 for μ, Eq. (1) for ρ, Eq. (7) for m/A_p, Eq. (2) for u, Eq. (4) for h and Eq. (3) for p. The explicit forms of $f(p)$ and $g(u)$ shown in Eq. (5) were used in the present solution, with $n=.69$ and $k=3.33\times10^{-5}$ sec/in. corresponding to JPN ballistite (a double base solventless propellant described in reference 1). A typical value of C is a value of .00556 computed from the burning rate data for JPN ballistite at 70°F as given in reference 1. The data on chemical

[4] See reference 1, pp. 39, 40.

[5] M. J. Zucrow, *Principles of Jet Propulsion and Gas Turbines* (John Wiley and Sons, Inc., New York, 1948), p. 105.

[6] A. M. Kuethe and J. D. Shetzer, *Foundations of Aerodynamics* (John Wiley and Sons, Inc., New York, 1950), pp. 134–139.

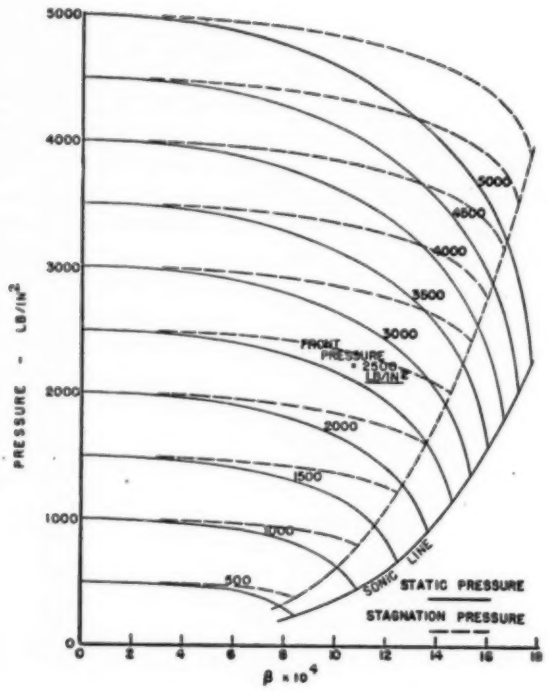


FIG. 2. Pressure distribution in a rocket motor.

equilibrium shown in Figs. 2.2 and 2.3 of reference 1 are also for a propellant of the JPN type composition. The summation in Eq. (7) involves repetition of the above computing sequence for successive increments $\Delta\beta$, the size of these increments being chosen to yield suitable accuracy.

Figures 2–4 show the relation of pressure, mass flux density m/A_p, and stream velocity to the parameter β for several values of the upstream pressure p_0, using the method of computation outlined above. Stagnation pressures were calculated by following isentropic paths in Fig. 2.3 of reference 1 to stagnation enthalpy. The results in Figs. 2–4 are applicable only to JPN ballistite propellant. The computations would have to be repeated for applications using propellants with different burning and thermodynamic characteristics.

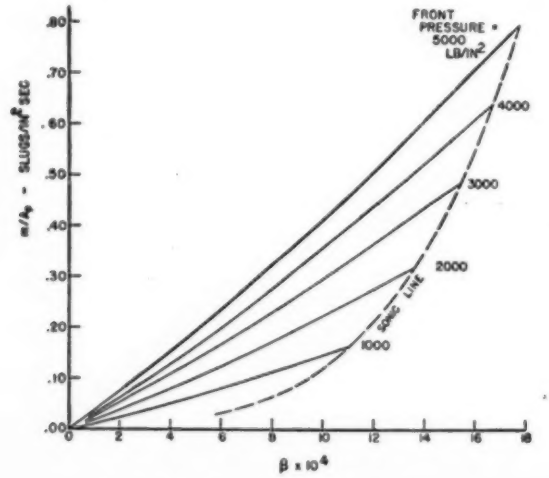


FIG. 3. Mass flux in a rocket motor.

SONIC FLOW

In performing the numerical calculations outlined above, it is found that the variables of state become progressively more sensitive to β as β is increased. As this sensitivity increases, the increments of β must be chosen smaller and smaller in order to obtain good accuracy, and the value of β consequently appears to converge to some limit.

It can be shown by an analytical treatment[7] of the problem using the assumption that γ and μ are constant, that for a given upstream pressure p_0, the maximum possible value of β is that for which sonic flow is reached

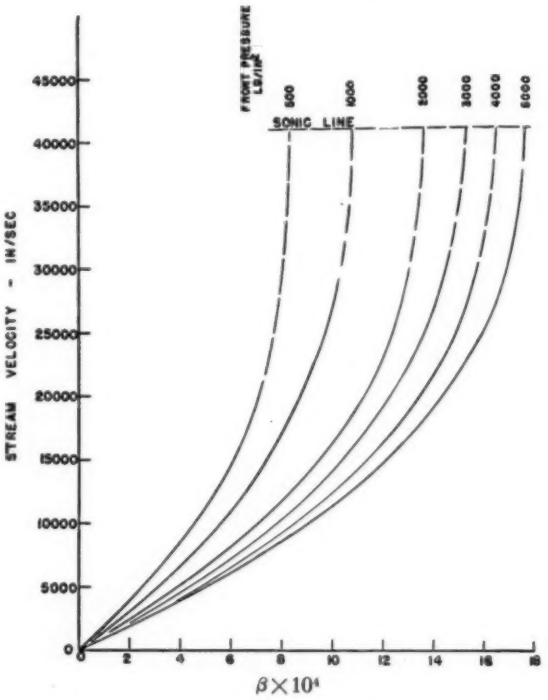


FIG. 4. Velocity distribution in a rocket motor.

at the exit of the channel. It can also be shown that the rate of change of velocity with β goes to infinity at the downstream end of the charge when sonic speed is reached there, and that the static pressure has the limit $p^*=p_0/(1+\gamma)$ where p^* is the pressure for sonic flow and γ is the ratio of specific heats for the gas.

This character of the dependency of the variables of state on the quantity β complicates calculation of the flow by numerical methods in the vicinity of β^*. In the present report the computed curves of pressure *vs* β were extrapolated to a pressure corresponding to $p^*=p_0/(\gamma+1)$, and the sonic limit drawn in at these points. A value for γ of 1.21 was used.

[7] Harold Grad, "Note on straight pipe jet motors," Comm. on Pure and Appl. Math., Vol. II, No. 1, March, 1949.

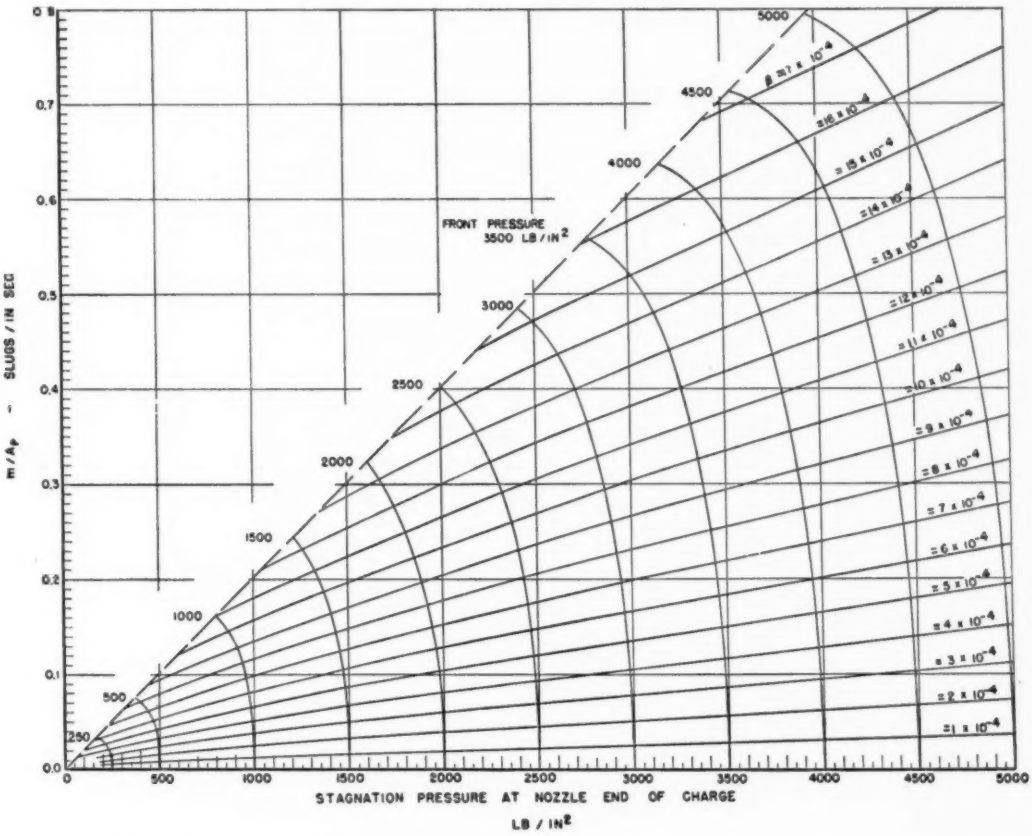


FIG. 5. Relation of mass flux density to pressure in a flow channel.

EQUILIBRIUM CONDITIONS

In a conventional rocket motor, the particular solution to the flow problem in the charge channel which actually prevails depends upon the size of the nozzle used downstream of the charge. This relation is established by the requirement that the mass burning rate of the charge equals the mass discharge rate of the nozzle. The discharge rate of the nozzle can be expressed satisfactorily by the relation[8,9]

$$m_d = C_D A_T p_s, \qquad (9)$$

where C_D is practically independent of A_T and p_s for a smooth contoured nozzle with sonic flow at the throat. A_T is the area of the nozzle throat, and p_s is the stagnation pressure at the nozzle entry.

By choosing particular values of β corresponding to particular propellant charges, curves can be plotted from Figs. 2 and 3 showing mass burning rate per unit port area *vs* downstream stagnation pressure. Such a family of curves is shown in Fig. 5. In the same figure, the discharge curve corresponding to Eq. (9) may be drawn, and the intersection of this curve with the curves for various values of β correspond to equilibrium conditions for the charges having those values of β and the nozzle corresponding to the discharge curve used. Figures 2–4 then provide the pressure, mass flux and velocity distributions.

In order to plot Eq. (9) in Fig. 5 it is necessary to divide the equation through by A_p. Thus the ratio $A_T/A_p = J$ becomes an important parameter in determining the equilibrium conditions for a given motor. Figure 6 shows the relation of the upstream pressure p_0 to the two parameters β and J, obtained by use of Figs. 5 and 2. This figure reduces the calculation of equilibrium pressure to the comparatively simple calculation of the parameters β and J, and in combination with Figs. 2, 3, and 4, establishes the pressure, mass flux, and velocity distributions in the motor. Within the assumptions made, this constitutes a solution to the problem of flow in the rocket motor.

RESULTS AND DISCUSSIONS

A computational procedure has been outlined for calculating the flow field in a solid propellant rocket with side burning charge. Solutions for the case of JPN ballistite propellant are presented. It is shown that the flow field depends upon a parameter β characteristic of the propellant charge geometry and a parameter J characteristic of the flow channel geometry. When the parameter $J \gtrsim 1$, the solution to the flow problem inside the motor depends only upon the value of β, while in more orthodox motor designs $J < 0.7$, and the solution to the flow problem depends upon the β and J.

[8] See reference 1, pp. 40, 52.
[9] See reference 3, p. 490.

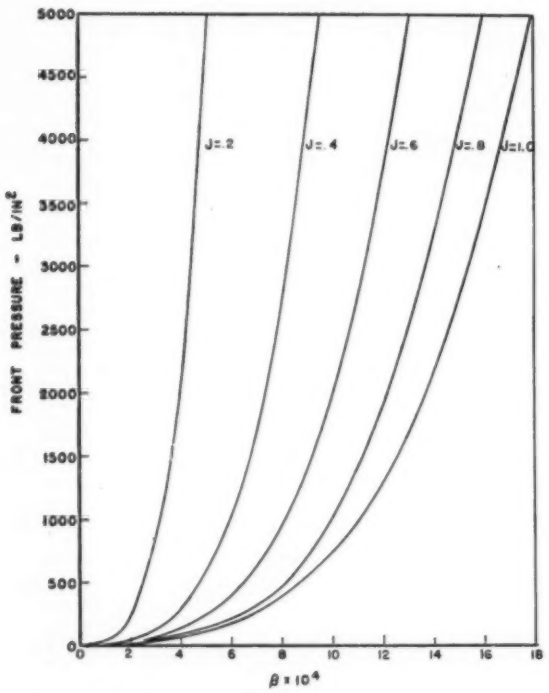


FIG. 6. Relation of equilibrium pressure to the charge parameter β and the channel parameter J.

The information yielded by the solutions presented here is required in the treatment of a number of practical problems in motor behavior such as sensitivity of equilibrium pressure to dimensional changes, propellant burning characteristics and propellant physical properties; relation of heat transfer to stream velocity, density and temperature; dependence of burning properties of propellants on adjacent stream conditions; and limiting and optimum conditions of motor and charge design.

NOMENCLATURE

A_p Cross-sectional area of the flow channel adjacent to the propellant charge.
A_T Cross-sectional area of the nozzle channel at its throat.
C Temperature factor in a burning rate rule for propellants (Eq. (5)).
C_D Parameter in a discharge rate rule for nozzles (Eq. (9)).
f Pressure factor in a burning rate rule for propellants (Eq. (5)).
g Stream velocity factor in a burning rate rule for propellants (Eq. (5)).
h Specific enthalpy of the propellant gas at the point x in the stream.
h_0 Specific enthalpy of the propellant gas at the upstream end of the flow channel (also the specific stagnation enthalpy of the stream elsewhere in the channel).
J Ratio of channel port area A_p to nozzle throat area A_T.
k Stream velocity parameter in a burning rate rule (Eq. (5)).
m Mass flux in the flow channel at the point x.
m_d Mass flux through the nozzle (discharge rate).
n Pressure exponent in a burning rate rule (Eq. (5)).
p Pressure at the point x.
p_0 Pressure at the upstream end of the flow channel.
p_s Stagnation pressure at the point x.
q Perimeter of a cross section contour of the burning surface at the point x.
r Burning rate normal to the surface of a solid propellant.
T Temperature of the propellant gas at the point x.
T_p Temperature of the solid propellant immediately prior to firing.
u Stream velocity.
x Position in the flow channel measured from the upstream or stagnation end.
γ Ratio of specific heats for the propellant gas.
β Ballistic parameter descriptive of charge geometry and propellant properties (Eq. (8)). May also be interpreted as position along the flow channel.
μ Average molecular weight of propellant gas.
ρ Density in the gas stream at the point x.
ρ_p Density of the solid propellant.
* Designates value of a quantity at a point where sonic flow is reached.

Direct Determination of the Flow Curves of Non-Newtonian Fluids

IRVIN M. KRIEGER AND SAMUEL H. MARON
Department of Chemistry and Chemical Engineering, Case Institute of Technology, Cleveland, Ohio
(Received July 25, 1951)

A method is proposed for the direct determination of the rate of shear-shearing stress curve of a fluid from data obtained in concentric cylinder viscometers. The method involves the use of a viscometer in which the ratio of outer to inner cylinder radii can be varied. It is further shown that to a close approximation two such radius ratios are sufficient to define the flow curve.

Data for two non-Newtonian systems are presented to show that the proposed method is valid.

INTRODUCTION

IN contrast to Newtonian fluids, whose viscosity coefficients are constant, many colloidal suspensions and macromolecular solutions exhibit variation of viscosity with rate of shear. The flow behavior of such "non-Newtonian" fluids is defined by the functional relationship between the velocity gradient (or rate of shear) dv/dy, and the shearing stress F, namely,

$$dv/dy=g(F). \tag{1}$$

This equation of flow is a differential equation which must be integrated in order to obtain relationships among the experimental variables measured in practical viscometers. For the two most common types of viscometer, the capillary and concentric cylinder systems, this integration can be readily performed if the function $g(F)$ is known. If, on the other hand, it is desired to determine $g(F)$ from the experimental variables, it becomes necessary to solve an integral equation.

Solution of the integral equation by differentiation has been accomplished for the capillary viscometer by Saal and Koens,[1] but to date no such solution has been obtained for the concentric cylinder viscometer. The present paper recasts the solution for the capillary viscometer into more convenient form, and presents a solution for the concentric cylinder viscometer in which the clearance between cup and bob is an independent variable. These solutions offer a consistent and convenient means for recovering the flow equations of non-Newtonian fluids from experimental data in the two common types of viscometer.

CAPILLARY VISCOMETER

Consider a fluid in laminar flow through a capillary of radius R and length L. The experimental observables are the rate of flow Q and the pressure drop P across the capillary. The shearing stress at a distance r from the axis of the capillary is

$$F=rP/2L, \tag{2}$$

while the velocity gradient is

$$dv/dy=-dv/dr. \tag{3}$$

Thus Eq. (1) becomes

$$-dv/dr=g(rP/2L).$$

The rate of flow is given by the integral

$$Q=\int_{r=R}^{0}\pi r^2dv=\int_0^R\pi r^2g\left(\frac{rP}{2L}\right)dr.$$

Changing the variable of integration from r to F, we have

$$Q=\frac{8\pi L^3}{P^3}\int_{F=0}^{F_m}F^2g(F)dF, \tag{4}$$

where

$$F_m=RP/2L \tag{5}$$

is the shearing stress at the capillary wall. An apparent fluidity may be defined by the relation

$$\phi_a=8LQ/\pi R^4P. \tag{6}$$

In the case of Newtonian flow ϕ_a is identical with the true fluidity, which is the reciprocal of the viscosity coefficient. Substituting (5) into (7),

$$\phi_a=\frac{4}{F_m{}^4}\int_0^{F_m}F^2g(F)dF. \tag{7}$$

It is significant that the equation relating ϕ_a and F_m for a capillary viscometer is independent of instrumental dimensions. Thus a graph of ϕ_a *vs* F_m for the same fluid in different capillaries should give a single curve. This theoretical prediction was first made by Reiner[2] on the basis of a power series for $g(F)$, and has been fully confirmed experimentally.

Equation (7) may be differentiated with respect to F_m and rearranged to give

$$\frac{g(F_m)}{F_m}=\phi_a+\frac{1}{4}\frac{d\phi_a}{d\ln F_m}. \tag{8}$$

This is essentially the result obtained by Saal and Koens. The quantity $d\phi_a/d\ln F_m=F_m d\phi_a/dF_m$ may be obtained from the slope of either a ϕ_a *vs* F_m or a ϕ_a *vs* $\ln F_m$ plot. Inaccuracy in its determination will reflect

[1] R. N. I. Saal and P. Koens, J. Inst. Petrol. Tech. **19**, 176 (1933).

[2] M. Reiner, *Deformation and Flow* (H. K. Lewis, London, 1949), pp. 108–110.

as a small error in $g(F_m)/F_m$, since ϕ_a is much larger than $d\phi_a/4d\ \ln F_m$.

Equation (8) offers then an accurate and convenient means for obtaining the flow curve of a fluid directly from capillary viscometer data.

CONCENTRIC CYLINDER VISCOMETER

Consider next a concentric cylinder viscometer with a bob of length L and radius R_1, and a cup of radius R_2. For definiteness assume that the cup rotates with angular velocity Ω, while a torque M exerted on the bob holds it stationary. Under laminar flow the velocity gradient at a distance r from the axis of rotation, where the fluid rotates with angular velocity ω, is

$$dv/dy=-rd\omega/dr, \tag{9}$$

while the shearing stress is

$$F=M/2\pi r^2L. \tag{10}$$

Equation (1) becomes in this case

$$-rd\omega/dr=g(M/2\pi r^2L). \tag{11}$$

Further, since $dr/r=-dF/2F$, Eq. (1) may be transformed to

$$d\omega=\tfrac{1}{2}g(F)dF/F,$$

which on integration between the bob, where $\omega=0$ and $F=F_1=M/2\pi R_1^2L$, and the cup, where $\omega=\Omega$ and $F=F_2=M/2\pi R_2^2L$, gives

$$\Omega=\frac{1}{2}\int_{F_1}^{F_2}\frac{g(F)}{F}dF. \tag{12}$$

If we let now

$$R_2^2/R_1^2=F_1/F_2=s^2, \tag{13}$$

Eq. (12) becomes

$$\Omega=\frac{1}{2}\int_{s^2F_2}^{F_2}\frac{g(F)}{F}dF. \tag{14}$$

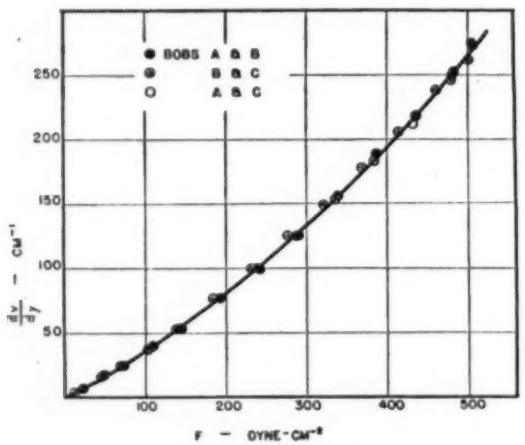


FIG. 1. Flow curve for a high solids GR-S latex by two-bob method.

Differentiation of Ω with respect to F_2 yields

$$\frac{d\Omega}{dF_2}=\frac{1}{2}\left(\frac{g(F_2)}{F_2}-\frac{s^2g(s^2F_2)}{s^2F_2}\right)=\frac{1}{2F_2}[g(F_2)-g(F_1)], \tag{15}$$

where instead of the desired quantity $g(F_2)$ or $g(F_1)$ the difference $g(F_2)-g(F_1)$ is recovered. This result was obtained by Mooney,[3] who used it as the starting point for an approximate solution.

While a general solution of Eq. (14) is not possible by this approach, a solution is possible for the special case of a finite bob rotating in a cup of infinite radius. For this case $F_2=0$, and Eq. (14) becomes

$$\Omega=\frac{1}{2}\int_{F_1}^{0}\frac{g(F)}{F}dF. \tag{16}$$

Differentiation of (16) with respect to F_1 gives

$$g(F_1)=-2d\Omega/d\ \ln F_1, \tag{17}$$

and hence $g(F_1)$ can be found from a plot of Ω *vs* $\ln F_1$.

A general solution of (14) can be obtained, however, by the simple expedient of considering the radius ratio s as the independent variable. To obtain Eq. (15), Eq. (14) was differentiated with respect to F_2 at constant s. Reversing the procedure and differentiating with respect to s at constant F_2, we get

$$(\partial\Omega/\partial s)_{F_2}=-g(s^2F_2)/s.$$

Rearranging, and recalling that $F_1=s^2F_2$,

$$g(F_1)=-s(\partial\Omega/\partial s)_{F_2}. \tag{18}$$

Equation (18) permits the evaluation of $g(F)$ from the variation of the angular velocity with radius ratio at a fixed value of shearing stress at the cup. A suitable apparatus for such a determination is a concentric cylinder system with several bobs of fixed length and varying radius. The condition of constant F_2 is equivalent to constant torque M. $(\partial\Omega/\partial s)_{F_2}$ can be determined from the experimental data by graphical or numerical differentiation. Equation (18) offers, therefore, an exact general solution of Eq. (16), whereby the flow equation of a non-Newtonian fluid may be recovered from experimental data obtained in a concentric cylinder viscometer.

TWO-BOB METHOD

A somewhat less exact but more convenient determination of $(\partial\Omega/\partial s)_{F_2}$ can be obtained with only two bobs of radii R_a and R_b, and hence radius ratios

$$s_a=R_2/R_a, \quad s_b=R_2/R_b.$$

With such a set-up is determinable the quantity

$$\left[\frac{\Omega_b-\Omega_a}{s_b-s_a}\right]_M=\left(\frac{\Delta\Omega}{\Delta s}\right)_{F_2}. \tag{19}$$

[3] M. Mooney, Rheology **2**, 210 (1931).

It is evident that $(\Delta\Omega/\Delta s)_{F_2}$ should furnish a reasonable approximation to the value of $(\partial\Omega/\partial s)_{F_2}$ at a mean radius ratio

$$s_m=\frac{s_a+s_b}{2}, \tag{20}$$

and hence Eq. (18) can be rewritten as

$$g(F_1)=-s_m(\Delta\Omega/\Delta s)_M. \tag{21}$$

The error E introduced by this approximation can be estimated by expanding Ω and $(\partial\Omega/\partial s)_{F_2}$ in Taylor's series in s about the point (Ω_a, s_a). It may be defined as

$$E=\frac{(\Omega_b-\Omega_a/s_b-s_a)_M-(\partial\Omega/\partial s)_{\text{at } s_m}}{(\partial\Omega/\partial s)_{\text{at } s_m}}, \tag{22}$$

which, on neglect of higher order differentials, gives

$$E=\frac{1}{24}(s_b-s_a)^2\frac{(\partial^3\Omega/\partial s^3)_{\text{at } s_m}}{(\partial\Omega/\partial s)_{\text{at } s_m}}. \tag{23}$$

The magnitude of E can be evaluated by using in Eq. (23) the exponential flow equation of Farrow, Lowe, and Neale,[4] namely,

$$g(F)=aF^N. \tag{24}$$

Here N is a measure of extent of non-Newtonian behavior. For Newtonian flow $N=1$, for "plastic" flow $N>1$, while for dilatant flow $N<1$. Substituting (24) into (14),

$$\Omega=\frac{1}{2}\int_{s^2F_2}^{F_2} aF^N\frac{dF}{F}=\frac{a}{2N}F_2^N(1-s^{2N}). \tag{25}$$

The partial derivatives of Ω may be gotten from (25) and substituted in (23) to give

$$E=\frac{(2N-1)(2N-2)}{24}\left(\frac{s_b}{s_a}-1\right)^2. \tag{26}$$

[4] F. D. Farrow, G. M. Lowe, and S. Neale, J. Textile Inst. **19**, T. 18 (1928).

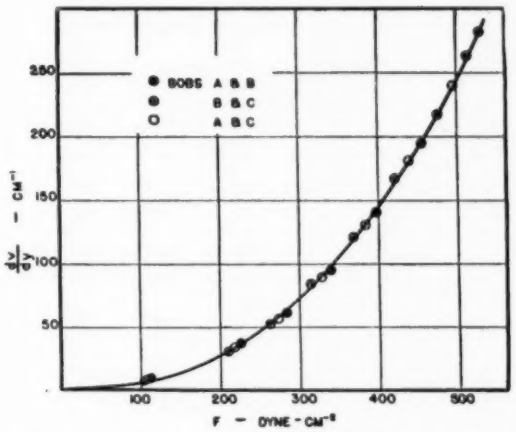


FIG. 2. Flow curve for a solution of polystyrene in methyl n-amyl ketone by two-bob method.

For Newtonian fluids E is zero, while even for liquids of large N the error may be kept within prescribed limits by suitable choice of $s_b/s_a=R_a/R_b$.

EXPERIMENTAL TESTS

The two-bob method described above has been tested and found to be convenient and accurate. Figures 1 and 2 show flow curves determined in this manner for a high solids synthetic rubber latex and a polystyrene solution. In both cases three different bobs were used; the three sets of points were obtained by taking the bobs pairwise. The fact that all three sets of points for the same system fall on a single curve confirms the validity of the proposed method.

ACKNOWLEDGMENT

The authors are indebted to the Office of Rubber Reserve, Reconstruction Finance Corporation, for sponsoring this research as part of the government Synthetic Rubber Program, and to Dr. Carl Moore for determining the data for Figs. 1 and 2.

Electron Diffraction Evidence for the Existence of Microstress in Evaporated Metal Films*

EBER K. HALTEMAN†
Department of Physics, Case Institute of Technology, Cleveland, Ohio
(Received August 20, 1951)

The line broadening in electron diffraction powder patterns from evaporated nickel films has been investigated. The pure diffraction broadening was found to vary with the Bragg angle in agreement with a microstress theory of broadening rather than with a particle size mechanism. The Fourier transform of the pure diffraction broadening has been determined and also found to be in agreement with a microstress mechanism.

MEASUREMENTS of the electrical resistivity of metals have shown that the resistivity depends upon impurities and imperfections. Microstrains will be associated with the imperfections in the lattice. These microstrains can be detected by the line broadening which they will produce in x-ray or electron diffraction powder patterns.

Line broadening in diffraction powder patterns may be due to two causes, small particle size or microstresses in the lattice. Broadening as a result of particle size alone is proportional to $\lambda \sec\theta$ while that caused by microstress alone is proportional to $E \tan\theta$ where E is the root mean square strain. An electron diffraction determination of line broadening, in which the Bragg angle θ is small, is well suited to distinguish between proportionality to $\lambda \sec\theta$ or to $E \tan\theta$.

EXPERIMENTAL

The substrata were prepared by evaporating silicon monoxide on a collodion film. The nickel films were prepared at a high rate of deposition and under less than 2×10^{-5} mm pressure so as to reduce contamination by foreign gas molecules. A series of transmission powder patterns was then taken in a high resolution double magnetic lens electron diffraction camera.

In microphotometering the plates, it was found that only an approximate linear relationship existed between the photographic density and the electron intensity. As the contour as well as the breadth of the broadened lines was desired, it was found necessary to experimentally determine the relationship between density and electron intensity for each plate.[1] The relative intensity was then corrected for the variation in atomic scattering factor.

A similar procedure was used to find the relative intensities resulting from the supporting substratum. This was then subtracted from the relative intensity for a sample of nickel upon a substratum to give the intensity caused by the nickel.

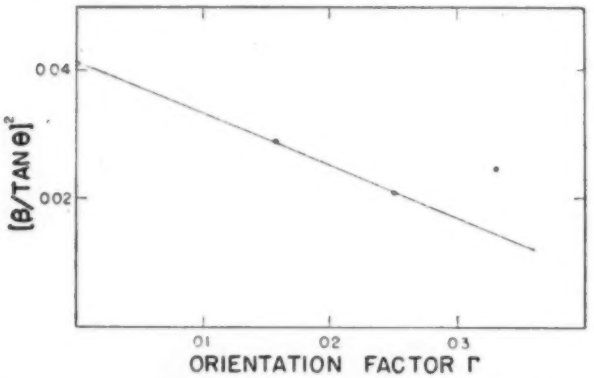


FIG. 1. $(\beta/\tan\theta)^2$ *vs* orientation factor γ, for evaporated nickel film.

RESULTS

The pure diffraction broadening β, i.e., that as a result of the effects of the sample alone, was determined by the Jones[2] method using the Warren[3] correction formula. Sodium chloride was used as a reference material which was assumed to consist of unstrained particles of large size.

Blackman[4] has shown that for any stress distribution which on the average acts symmetrically with respect to the crystallographic axis the root mean square strain E will be given by

$$E^2 = A - B\gamma, \tag{1}$$

where A and B are constants and γ is the orientation factor. As the pure diffraction broadening β is proportional to $E\tan\theta$,

$$(\beta/\tan\theta)^2 \sim A - B\gamma. \tag{2}$$

Typical results for a 50A nickel film are shown in Fig. 1. A straight line with a negative slope is obtained in agreement with the microstress mechanism for line broadening. The value of β for the (111) lines $\gamma=0.333$ was uniformly high for all samples. No reason is known for this. If a stacking fault occurred in the (111) plane, the broadening would be only $\frac{3}{4}$ as great as that for a random distortion. Thus, if some stacking faulting does occur, one would expect the value of β for the (111) planes to be somewhat less than for an isotropic stress alone.

* Part of a thesis submitted to the Faculty of the Graduate School of Case Institute of Technology in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

† Now at Westinghouse Atomic Powder Division, Bettis Field, Pennsylvania.

[1] G. P. Thomson, and W. Chochrane, *Theory and Practice of Electron Diffraction* (Macmillan and Company, Ltd., London, England, 1939), p. 244.
[2] F. W. Jones, Proc. Roy. Soc. **166A**, 16, (1938).
[3] B. E. Warren, J. Appl. Phys. **12**, 375 (1941).
[4] N. M. Blackman, Phys. Rev. **70**, 69 (1946).

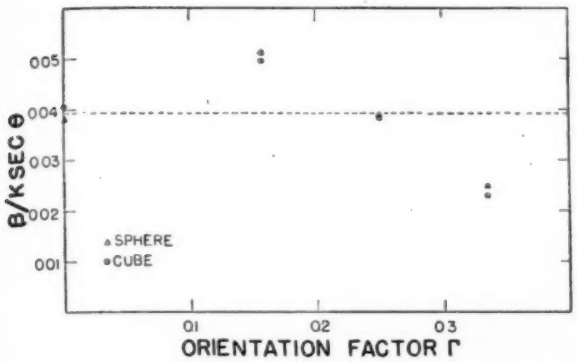


FIG. 2. (β/K secθ) *vs* orientation factor γ, for evaporated nickel film. Dashed line indicates expected curve for a particle size mechanism.

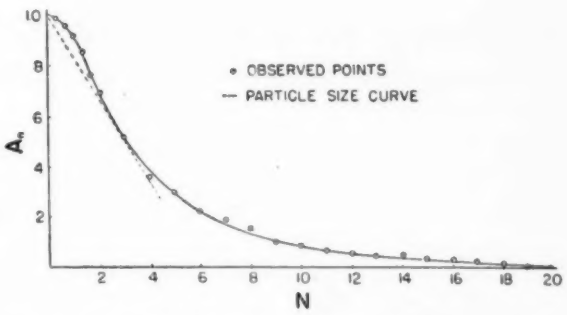


FIG. 3. A_n *vs* n for the (111) reflection of an evaporated nickel film. (a) Observed point. (b) Dashed line indicates expected curve for a particle size mechanism.

The pure diffraction broadening caused by a particle size mechanism is given by the equation

$$\beta = (K\lambda \sec\theta)/L, \tag{3}$$

where K is the Scherrer "constant" and L an average dimension of the particle. The value of K depends upon the external shape of the diffracting particle and on the order of the spectrum, but in all cases its value is nearly unity. The value of β/K secθ should then be constant and independent of the orientation factor γ. The results are shown in Fig. 2 which indicates that β/K secθ is not independent of γ and that the diffraction broadening is not a result of a particle size mechanism.

Warren and Averbach have shown that from a knowledge of the Fourier coefficients of the pure diffraction broadening β, it is also possible to distinguish between a particle size mechanism and a microstress mechanism.[5] The Fourier transform of the pure diffraction broadening caused by a particle size mechanism alone is in the form of a triangle with a toe. The toe would be a result of a distribution of different particle sizes in a polycrystalline material. The Fourier transform of the pure diffraction broadening because of a microstress mechanism is in the form of an error curve. The transform of the pure diffraction broadening β of the (111) reflection for a thin nickel film is shown in Fig. 3. This result indicates that the broadening is caused by a microstress mechanism.

Knowing the Fourier coefficients it is possible to obtain the distortion ΔL as a function of the distance L from the center of the distortion.[6] In the (111) planes of an unstrained nickel lattice the atoms have a spacing of 2.03×10^{-8} cm. At this distance in the [111] direction from a distortion the microstrain, $\Delta L/L$, was found by using the method of Warren and Averbach to be approximately 0.05. These values of strain exist only over short distances in the crystal lattice and indicate the localized nature of the microstrain.

Values of strain of about 0.05 can be explained on the basis of distortions which could be produced when stacking faults appear in the (111) plane. The face centered cubic structure of nickel is formed by stacking layers of atoms in the arrangement abcabc···. The hexagonal, close packed structure is formed by stacking layers in the arrangement ababab···. A stacking fault will arise when the proper sequence of layers is not maintained.

A simple form of distortion in stacking faulting would occur when the second layer of nickel atoms was arranged in the form of a "*b*" plane at one place and in the form of a "*c*" plane at another place. In the region where the "*b*" and "*c*" arrangements would attempt to join, a small gap would form. If the last correctly placed "*b*" and "*c*" atoms were arranged in straight parallel lines and another atom was squeezed into the gap between, the strain would be approximately 0.105. This is computed on the basis of stacking inelastic spheres and dividing the strain equally on either side of the (111) plane. As atoms are not inelastic spheres, one would expect that several adjacent layers would be effected by the inserted atom. Thus the strain of the nearest neighbors would be somewhat less than 0.105.

One could postulate other ways in which a distortion could arise in stacking atoms. In no case does the initial strain have a value less than 0.10 when the atoms are simply considered as hard spheres. Thus the observed initial strain of 0.05, although many times the yield strain of bulk nickel, does not appear to be unreasonable.

CONCLUSION

For the nickel films analyzed the line broadening was primarily a result of a microstress mechanism rather than a particle size mechanism. These microstresses were very large and exist only over short distances in the crystal lattice. Distortions caused by stacking faults are a possible explanation for the large microstrains.

ACKNOWLEDGMENT

The author wishes to express his thanks to Dr. E. C. Crittenden, Jr., for his suggestion of this work and for his assistance in its completion.

[5] B. E. Warren, and B. L. Averbach, J. Appl. Phys. **21**, 595 (1950).

[6] B. E. Warren, and B. L. Averbach, *ibid.*

Letters to the Editor

The Size of Latex Particles by X-Ray Scattering*

B. R. Leonard, Jr., J. W. Anderegg, Paul Kaesberg,† and W. W. Beeman
Department of Physics, University of Wisconsin, Madison, Wisconsin
(Received October 17, 1951)

THE discovery of the extremely uniform size of the spherical particles of polystyrene Dow latex 580 G, lot 3584 has caused considerable interest in the possible use of these particles as a calibration standard for electron microscopes.[1] A survey of determinations of the size of these particles has been presented by Gerould.[2] Determinations which depend on a measurement with the electron microscope have led to an accepted value of 2590A for the particle diameter. Kern and Kern,[3] however, have reported a value of 2590A for the diameter of the unshadowed particles but a value of 2840A for obliquely shadowed particles. Recently, investigators using other methods have obtained diameters between these two. Dandliker[4] has reported a value of 2720A as determined by light scattering and Yudowitch[5] has reported a value of 2740A by small angle x-ray scattering. Our result is in close agreement with these later values.

We have used the method of small angle x-ray scattering[6] to determine the particle diameter from the secondary diffracted maxima characteristic of spheres. The experimental arrangement used was similar to that of Ritland, Kaesberg, and Beeman.

The angular distribution of scattered intensity from spherical particles of uniform electron density is given by the well-known function

$$I(\theta) \propto MN^2\left\{\frac{\sin(kR)-(kR)\cos(kR)}{(kR)^3}\right\}^2,$$

where M is the number of irradiated spheres, N is the number of electrons per sphere, $k=4\pi\dfrac{\sin\theta}{\lambda}$, θ is one-half the angle of scattering, λ is the x-ray wavelength, and R is the particle radius. The positions of the secondary maxima and minima of this function are known and are used to determine the particle radius.

The scattered x-ray intensity depends on the square of the difference of the electron density between the particle and the surrounding medium. Since the electron density of the latex is very nearly that of water, the scattering from a suspension of latex in water is so small that its measurement is not feasible. Hence our samples were run dry or in a suspension of methyl alcohol. The samples were dried over phosphorous pentoxide at atmospheric pressure, and the wet samples were obtained by resuspending in methyl alcohol.

The accompanying figure shows representative curves of the scattered intensity vs angle obtained from two different thicknesses of dry sample and a wet sample. The curves all show the smearing effect resulting from finite slit dimensions. The thick sample curve also shows the effect of multiple scattering in smearing the peaks. We have observed the first seventeen maxima from the thick sample and the first eight maxima from the thinner samples and the wet samples. The accuracy with which the particle diameter can be determined is directly proportional to the scattering angle, and in the analysis of the data the values obtained are weighted in proportion to the angle. A weighted average of 133 separate values of maxima and minima yields 2731A for the particle diameter. The weighted average deviation of the individual measurements from the mean is 34A, and the statistical accuracy of the mean is about ±3A. The results from the wet and dry samples are in agreement. We have also measured a sample of the latex used by Yudowitch in his determination and this gives a result consistent with our value.

* Research supported in part by the ONR and in part by the Wisconsin Alumni Research Foundation.
† Department of Biometry and Physics, University of Wisconsin.
[1] R. C. Backus and R. C. Williams, J. Appl. Phys. **19**, 1186 (1948); R. C. Backus and R. C. Williams, J. Appl. Phys. **20**, 224 (1949); G. D. Scott, J. Appl. Phys. **20**, 417 (1949).
[2] C. H. Gerould, J. Appl. Phys. **21**, 183 (1950).
[3] S. F. Kern and R. A. Kern, J. Appl. Phys. **21**, 705 (1950).
[4] W. B. Dandliker, J. Am. Chem. Soc. **72**, 5110 (1950).
[5] K. L. Yudowitch, J. Appl. Phys. **22**, 214 (1951). The value reported by Dr. Yudowtich in this article was 2780A. However, Dr. Yudowitch informed the authors by private communication prior to this experiment that he had revised his result to 2740A.
[6] Ritland, Kaesberg, and Beeman, J. Chem. Phys. **18**, 1237 (1950); A. Guinier, Ann. phys. **12**, 161 (1939); O. Kratky, J. Polymer Sci. **3**, 195 (1948).

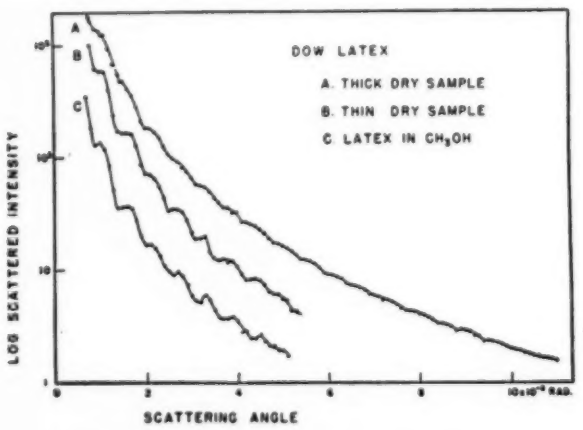


FIG. 1. Representative curves of scattered x-ray intensity as a function of angle for polystyrene Dow latex.

The Effect of Pressure on Scintillation Phosphors*

A. J. Reinsch and H. G. Drickamer
Department of Chemistry, University of Illinois, Urbana, Illinois
(Received October 19, 1951)

IN the course of development of high pressure tracer techniques, a study has been made of the pressure coefficient of cadmium tungstate and anthracene phosphors in the range from atmospheric pressure to 10,000 atmospheres. The windows used were synthetic sapphire with the "c" axis oriented perpendicular to the flats.

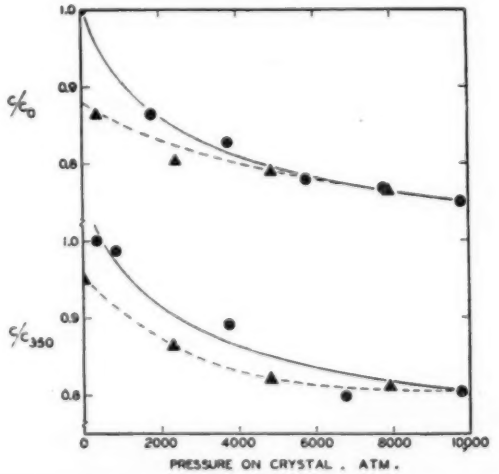


FIG. 1. The effect of pressure on counting rate for a $CdWO_4$ crystal on a sapphire window.

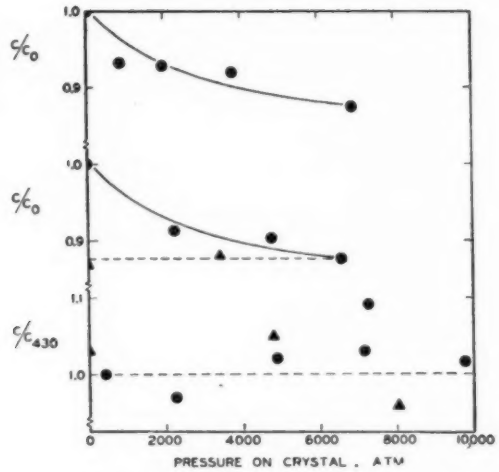


FIG. 2. The effect of pressure on counting rate for an anthracene crystal on a sapphire window.

The windows were one-half inch in diameter and one-half inch thick, and covered a quarter-inch opening. The sapphire and the metal were ground and lapped flat, and a Poulter[1] type seal was used. The bomb design and the other seals were standard of the type commonly used by P. W. Bridgman. In one test a plate glass window of approximately the same dimensions was used.

In each case the crystal was fastened to the window with Canada Balsam, solid sulfur containing S^{35} was piled on the crystal, and covered with dissolved Lucite to protect it from the pressure-transmitting fluid (heptane). The Lucite was permitted to harden over night. The light was conducted from the window to a photomultiplier tube by a Lucite rod. The tube was cooled to dry ice temperature. The tube circuit and general setup was only slightly modified from that previously described.[2]

The pressure was measured by a manganin gauge, calibrated from the freezing point of mercury at 0°C (7394 atmos).

The results for cadmium tungstate are shown in Fig. 1, those for anthracene in Fig. 2. In each case there was an irreversible decrease in counting rate, larger on the first run than on later runs, due to the fouling of the window. This effect was not large enough to mask the distinct differences in the behavior of the crystals. In the case of cadmium tungstate there was a reversible pressure effect on the scintillation efficiency, amounting to 15 to 20 percent decrease in efficiency at 10,000 atmospheres. For anthracene, there was no noticeable effect of pressure except on the fouling of the windows.

An explanation of the phenomena can be obtained from Fig. 3, which shows the absorption curve for cadmium tungstate and the response curve of the photomultiplier tube. The strongest emission band for $CdWO_4$ is at 5200A, that for anthracene is at 4400A. From the absorption curve for $CdWO_4$, it is clear that no reasonable shift of the emission or absorption band combined would alone be sufficient to give the measured effect. However, if the emission band for $CdWO_4$ would shift 100A toward higher wavelengths, the relative response of the photomultiplier tube would decrease by 14 percent. A similar shift in the emission band for anthracene would result in a 3 percent decrease in relative tube response.

The authors are indebted to Dr. R. J. Maurer and his colleagues of the Physics Department of the University of Illinois for running the absorption curve on $CdWO_4$, and for helpful conversations concerning the interpretation of the results.

* This work was supported in part by the AEC.

[1] T. C. Poulter, Phys. Rev. **35**, 297 (1930).
[2] Timmerhaus, Giller, Duffield, and Drickamer, Nucleonics **6**, 37 (1950).

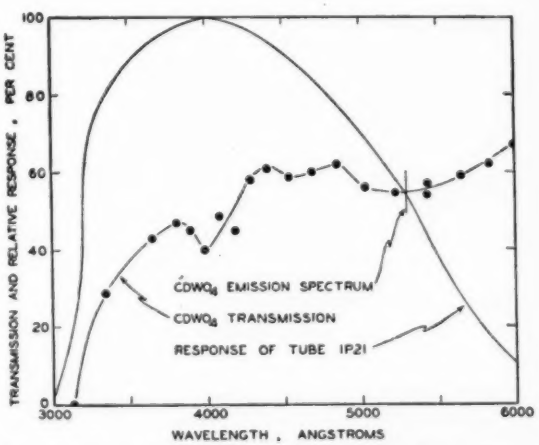


FIG. 3. $CdWO_4$ emission and transmission characteristics compared with response of photomultiplier tube [*P2*].

Direct Printing of Shadowed Electron Micrographs

R. M. FISHER
Research Laboratory, United States Steel Company, Kearny, New Jersey
(Received November 5, 1951)

SHADOWCASTING[1] of electron microscope specimens is one of the most useful and often used techniques of electron microscopy. As is well known, shadowing, besides giving a three-dimensional aspect to micrographs, improves the contrast of organic replicas and very fine particulate matter.

Unfortunately, to obtain "black" shadows on prints a photographic reversal is required. This is done by making a positive transparency from the original negative and then making prints from the positive. This extra step increases the chance of scratches, pin-holes, and loss of sharpness in the print and takes considerable time, especially when many routine pictures are taken. For these reasons many laboratories make direct prints from the negatives and as a result the shadows on such prints are white. The "shadow cap" on heavily shadowed replicas is often mistaken for the shadow, leading to confusion as to the relief of detail on the surface.

FIG. 1. Cementite lamellae extracted from steel. Direct printed on reversing paper. 16,000X.

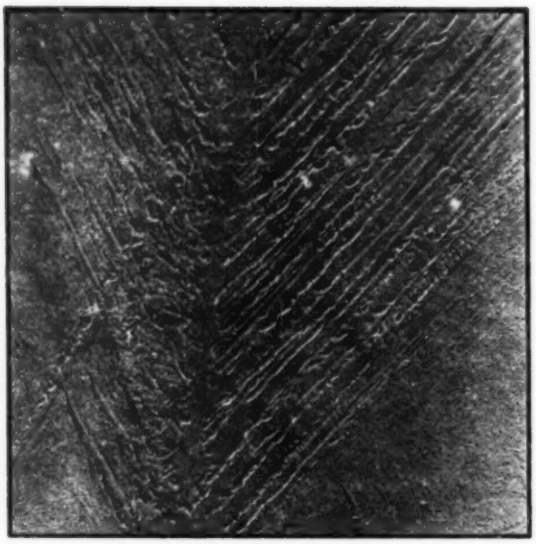


FIG. 2. Replica of bainite region along grain boundary in steel. Direct printed on reversing paper. 4000X.

A photographic method is available to make negative prints directly with no more effort than is required to make positive prints. By using Kodak reversing paper "kind 1198" and Kodak "special developer 426" reverse prints can be made directly, requiring no more steps than the usual positive print processing. Exposure and development times are comparable to the normal procedure. The "kind 1198" paper is a unicontrast paper, but contrast can be varied from about F1 to F3 by variation in the amount of sodium hydroxide added to the developing solution.

Figures 1 and 2 are examples of this direct printing method. Figure 1 is a uranium shadowed electron micrograph of lamellae of cementite extracted from steel and mounted on collodion. Figure 2 is a uranium shadowed replica of a 1048 steel showing bainite along a grain boundary.

I wish to thank John Faber and William Swann of the Eastman Kodak Company for supplying this paper and developer to me before it became generally available.

[1] R. C. Williams and R. W. G. Wyckoff, J. Appl. Phys. **15**, 712 (1944).

Dielectric Constants of Various Gases at 9470 Mc*

CLAYTON M. ZIEMAN
Wabash College, Crawfordsville, Indiana
(Received August 22, 1951)

IN determining the dielectric constants of various gases at 9470 Mc (wavelength ≈3 cm), use was made of the principle that the resonance frequency of a cavity depends upon the dielectric constant and the permeability of the substance filling it. The apparatus consisted of a frequency-modulated reflex klystron transmitting energy to two transmission-type cavities by means of a branched wave-guide system. One of the cavities was enclosed in a brass chamber with a window in one face to permit observation of the instruments in it. The gas under investigation was first chemically dried and then passed into the evacuated cavity through a small hole in the center of the broad side of the connecting wave guide. The wave-guide branches were terminated by crystal detectors whose outputs were separately amplified. The amplified outputs were fed to an electronic switching circuit and then presented on an oscillograph screen. The modulating voltage served as the oscilloscope sweep.

As the chamber containing the test cavity was gradually evacuated, the resonance pulse for this cavity moved along the time-axis. With the chamber evacuated, two pulses were placed in coincidence. As the gas under investigation was admitted to the chamber, the pulse for the test cavity moved back along the time-axis. The frequency change was read from the micrometer screw used to tune the reference cavity after the two pulses were again placed in coincidence. To increase the accuracy, a vernier and gear system were added. A rotation of 14° of the tuning screw gave a rotation of 1535° of the knob driving the system.

In Table I are listed the dielectric constants for dry air, CO_2, N_2, and O_2, all reduced to S.T.P. These results are the averages of 45 measurements made with dry air, 12 with CO_2, 16 with N_2, and 38 with O_2. The dominant contribution to the values of the precision measures resulted from the residual error associated with the change of frequency per degree of rotation of the gear system driving the reference cavity plunger. The gases were obtained in cylinders from the Matheson Company, Joliet, Illinois, and were claimed to be about 99 percent pure.

TABLE I. Dielectric constants at 9470 Mc by method of electronic switching.

Gas	$(\epsilon_0-1)\times10^4$
Dry Air[a]	5.754±0.021
N_2	5.870±0.020
O_2[b]	5.320±0.021
CO_2	9.875±0.041

[a] Corrected for permeability effect, $\mu_0-1=0.004\times10^{-4}$.
[b] Corrected for permeability effect, $\mu_0-1=0.018\times10^{-4}$.

The foregoing apparatus was next converted into a microwave refractometer[1] by adding to it a pulse sharpener and a meter circuit. This refractometer, calibrated each time with dry nitrogen, was used to make the additional measurements summarized in Table II.

TABLE II. Dielectric constant at 9470 Mc by the microwave refractometer.

Gas	$(\epsilon_0-1)\times10^4$
Dry Air	5.74±0.02
H_2	3.55±0.05
O_2[a]	5.31±0.03

[a] Corrected for the permeability effect, $\mu_0-1=0.02\times10^{-4}$.

Twenty-five measurements were made for each gas.

The foregoing work was begun and largely completed without knowledge of a similar project carried out by a somewhat different method at the National Bureau of Standards by Birnbaum, Kryder, and Lyons.[2]

The support of the Research Corporation in carrying out the described project is gratefully acknowledged.

* Clayton M. Zieman, Phys. Rev. **83**, 243(A) (1951).
[1] George Birnbaum, Rev. Sci. Instr. **21**, 169 (1950).
[2] Birnbaum, Kryder, and Lyons, J. Appl. Phys. **22**, 95 (1951).

A Physical Theory of Rubber Reinforcement

A. M. BUECHE
General Electric Research Laboratory, Schenectady, New York
(Received October 15, 1951)

THE stiffening produced by the introduction of solid fillers into elastomers has been treated in a different and more detailed way than has been used in previous theories.[1-3] It is the purpose of this note to outline briefly the method of treatment

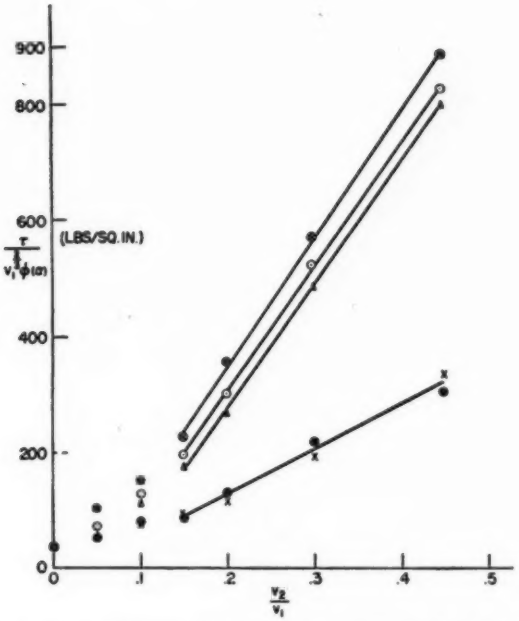


FIG. 1. Tension at 300 percent elongation for Neoprene-carbon black mixtures. The blacks are ●-MT, X-FT, ⊙-SRF, △-HMF, ⊗-MPC.

and to compare the results with some pertinent data in the literature.

The gum sample was considered to consist of N linear polymer molecules of mass m. To this was added filler particles to which the polymer could adhere. It was assumed that the number, n, points of adherence was proportional to the filler surface and for the case of spherical particles of radius r, was $n=4\pi r^2 f$ per particle. The quantity f is then the number of points of adherence per unit area of filler.

Making the assumptions that the points of adherence are added at random and that a molecule will contribute to the elasticity only when it has two or more points of adherence, both of which seem reasonable except at very low filler concentrations, it can be shown that the number of polymer segments which can contribute to the elasticity (act like molecular springs) is

$$\nu=N_0\left\{\left(1-\frac{1}{N}\right)^n-\left(1-\frac{n}{N}\right)\right\}. \tag{1}$$

This predicts that the first few filler particles will be relatively ineffective and that as more are added, the number of molecular springs approaches

$$\nu=n-N. \tag{2}$$

For simplicity Eq. (2) rather than Eq. (1) was used in the further development. Because of this it is to be expected that the final result will deviate from experimental data at low filler concentrations.

The usual theories[4] of rubber elasticity were then modified to include the fact that in a filled sample, the volume fraction of polymer, v_1 is not unity. This led to the result that

$$v_1^{\frac{1}{3}}\tau=kT(\nu/V)\varphi(\alpha), \tag{3}$$

where k is Boltzman's constant, T is the absolute temperature, V is the volume of the filled sample, τ is the tension (force per unit initial cross-sectional area), and

$$\varphi(\alpha)=\alpha-1/\alpha^2,$$

where

$$\alpha=1+\Delta L/L_0 v_1^{\frac{1}{3}}.$$

Here ΔL is the difference between the initial length, L_0, and the length at which the tension is measured. By a combination of Eqs. (2) and (3) and the realization that $c/2$ cross-links are added per unit volume of gum by the vulcanization process, it was shown that

$$\tau/\varphi(\alpha)v_1^{\frac{1}{3}}=kT[c-\rho/m+(3f/r)(v_2/v_1)]. \tag{4}$$

where ρ is the density of the pure polymer and v_2 the volume fraction of filler.

It is interesting to notice that the magnitude of the filler effect is determined by the quotient f/r. Because of the difficulty of producing small particles of different sizes but identical surface characteristics (same f) and at the same time keeping c constant it is little wonder that there are so many conflicting statements in the literature.

Catton and Thompson[5] have investigated the system carbon black-Neoprene under conditions where there is evidence that the presence of carbon does not influence the vulcanization process (c constant). Their data are plotted as suggested by Eq. (4) in Fig. 1. As predicted the data give a linear plot at higher filler concentrations and the expected curvature at low concentrations.

The surface area, $1/f$, necessary for one point of adherence as calculated from the slopes of the linear portions and electron microscopic particle sizes is given in Table I.

TABLE I.

Carbon	$2r^a(m\mu)$	$(1/f)A^2$
MT	274	1.74
FT	75	6.36
SRF	83	2.09
HMF	52	3.35
MPC	28	5.97

[a] Rubber Age 55, 469 (1944).

The areas obtained fluctuate somewhat and seem a little too small but are of the expected order of magnitude. The assumption of perfectly round filler particles and the neglect of molecular intertwining would all introduce errors in this direction.

The foregoing methods can be used to estimate the swelling of filled polymers. It is planned to include this in a more detailed publication.

[1] J. Rehner, Jr., J. Appl. Phys. 14, 638 (1943).
[2] H. M. Smallwood, J. Appl. Phys. 15, 758 (1944).
[3] E. Guth, J. Appl. Phys. 16, 20 (1945).
[4] For a review see P. J. Flory, Chem. Rev. 35, 51 (1944).
[5] N. Catton and D. Thompson, Rubber Chem. and Tech. 22, 812 (1949).

Announcements

New Vice-President at University of Pennsylvania

DR. Charles W. MacGregor, professor of applied mechanics and head of the materials division at Massachusetts Institute of Technology, has been named vice-president of the University of Pennsylvania.

In the office at Pennsylvania, newly-created by the trustees, Dr. MacGregor will have jurisdiction over the engineering and scientific departments of the university. His administration will also include the Towne Scientific School and the Moore School of Electrical Engineering. Dr. MacGregor will begin his work at Pennsylvania in February.

National Science Board Elects Barnard, Chairman, and Fred, Vice-Chairman at Second Annual Meeting

CHESTER I. Barnard, president of the Rockefeller Foundation, succeeds James B. Conant as Chairman of the National Science Board of the National Science Foundation. Edwin B. Fred, president of the University of Wisconsin, was re-elected vice-chairman of the National Science Board. The new chairman and vice-chairman were elected for two-year terms of office as prescribed by the National Science Foundation Act of 1950. The annual meeting marks the beginning of the second year of operation for the National Science Foundation, which was created by Congress early in 1950.

Four members of the Executive Committee, whose terms had expired, were also re-elected. These were as follows: Mr. Barnard; Detlev Bronk, president of the Johns Hopkins University and president of the National Academy of Sciences; Lee A. DuBridge, president of the California Institute of Technology; and Elvin C. Stakman, chief of the Division of Plant Pathology and Botany at the University of Minnesota.

Standards on Letter Symbols for Hydraulics

THE American Standards Association Subcommittee Y10.2 (formerly, Z10.2), Letter Symbols for Hydraulics, has recently been reconstituted for the purpose of reviewing and revising the 1942 Standards on Letter Symbols for Hydraulics.

The committee would appreciate receiving comments from interested parties concerning the 1942 Standards and any suggested additions or modifications thereof. It has been tentatively decided to consider the field of hydraulics as being concerned with all phases of the motion of liquids. An attempt will be made to include symbols in the field of hydrology and water-hammer. Suggestions and comments should be directed to the Chairman, Dr. J. M. Robertson, P. O. Box 30, State College, Pennsylvania.

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 1 JANUARY, 1952

PROGRAM

Proceedings of the Electron Microscope Society of America

THE annual meeting of the Electron Microscope Society of America was held at the Franklin Institute in Philadelphia, Pennsylvania, November 8, 9, and 10, 1951. Titles and abstracts of the papers presented are given below. The name of the person presenting the paper is marked with an asterisk.

General Subjects

CHAIRMAN: R. D. HEIDENREICH

1. An Empirical Treatment of the Influence of Morphology and Accelerating Voltage on Electron Diffraction Patterns. S. G. ELLIS, *RCA Laboratories Division, Princeton, New Jersey.*—Using films of known thickness, and accelerating voltages of 50, 100, and 150 kv to produce diffraction patterns, it has been possible to construct an empirical theory of plural electron scattering. This theory is used to calculate the contrast (ring intensity/background intensity) for specimens with different morphologies. While the contrast usually increases rapidly with voltage, an increase in voltage in the range used does not greatly increase the maximum permissible specimen thickness.

2. An Intermediate and Diffraction Lens for the RCA Electron Microscope Type EMU. BENJAMIN M. SIEGEL, *Department of Engineering Physics, Cornell University, Ithaca, New York.*—An intermediate lens has been built for the RCA EMU similar to the one described by Hillier[1] but with a removable aperture. The projector lens circuit has been rewired so that the coil of the intermediate lens can be switched in series with the projector coil as well as in parallel. The higher current available makes it possible to focus the image focal plane of the objective lens at the object distance of the projector lens and thus obtain an electron diffraction pattern from the specimen *in situ*. An aperture 300 microns in diameter will give the diffraction pattern from the crystalline material contained on an area about 3 microns in diameter on the specimen.

[1] J. Hillier, J. Appl. Phys. 21, 785 (1950).

3. Electron Diffraction from Small Areas of Electron Microscope Specimens Using an RCA Electron Microscope. J. D. BOADWAY, *Shawinigan Chemicals, Ltd., Shawinigan Falls, Quebec, Canada.*—When a fine aperture is used in the condenser of an RCA electron microscope, diffraction patterns may be taken of a specimen when in the position above the objective which is normally used for direct electron photography. Quite small areas may be studied and by transition of focal conditions the diffraction pattern may be related to a

micrograph of the same area. Examples are shown of diffraction from specimens of mica and NaCl and also from differing crystalline components of lime hydrate produced on slaking calcium carbide.

4. An Objective Aperture Alignment Device. MARK E. GETTNER, *The Sloan-Kettering Institute for Cancer Research, New York, New York.*—In objective aperture alignment, ability to see the motion of the aperture on the final viewing screen and application of the stage control knobs to this purpose have proven of value. The device described consists of an aperture-holding ring, a friction platform on which the aperture holder rests, and a probe similar to a specimen holder resting in the stage and making contact with the aperture holder. In operation, the aperture is placed into the aperture holder, which is then put onto the platform. The top of the objective pole piece is screwed on and the pole piece inserted into the microscope. The stage is placed in its central position, and the probe is dropped into it. After evacuation, the stage is moved about carrying with it the probe and aperture holder until the aperture appears on the final viewing screen. The fine movement of the stage screws is used to center the aperture precisely. The probe is replaced by the specimen holder, and the microscope is in position for use. The entire centering procedure takes from five to ten minutes.

5. The Reliability of Internal Standards for Calibrating Electron Microscopes. JOHN H. L. WATSON,* *Edsel B. Ford Institute for Medical Research, Henry Ford Hospital, Detroit, Michigan,* AND WILLIAM L. GRUBE, *Research Laboratories Division, General Motors Corporation, Detroit 2, Michigan.*—Conclusions important to the practice of electron microscopy are drawn from experiments. (1) Particles of Dow Latex 580G are unstable under electron beams and best micrographed at low intensity. (2) Although uniform and spherical the mean diameter of 580G spheres is not dependable, being 3000A in the present work. (3) The 580G is improved as an internal standard if calibrated by some other suitable standard (not the reverse) each time before use. (4) Fresh replicas of diffraction gratings are still the most practical standards, reproducing gratings to two percent. (5) By an independent optical method (spectroscopy), one can measure the grating space of a replica when it is on an eighth inch, 200-mesh, specimen screen under exact calibration conditions for electron microscopes. (6) Specific replicas, and from them specific preparations of 580G, are checked easily and accurately by the spectroscopic method. Replica age or other factors which may affect accurate reproduction of an original grating are then no longer important in calibration.

6. Experimental Improvements in the Performance of an Electron Microscope. JAMES HILLIER, *RCA Laboratories Division, Princeton, New Jersey.*—The continuing exploration of the factors which prevent the consistent attainment of maximum resolution has led to a number of promising experimental improvements in the magnetic electron microscope and to the identification of at least one previously ignored difficulty. The modification of the electron gun to give nearly an order of magnitude increase in intensity showed that space charge diffusion of the images sets a definite upper limit on the total bombardment of the specimen. The use of a small, centerable condenser aperture with this gun eliminated the space charge and greatly reduced the thermal loading of the specimen. These developments and the use of a low power stereoscopic microscope to examine the final image have made it possible to observe visually image defects down to the 30A level and to compensate objectives by visual observation. The sensitivity of the instrument to stray magnetic fields and the lack of coincidence between voltage and magnetic centers led to the use of a double objective in which the difficulties are considerably reduced and which can provide electronic magnifications up to 100,000. While this system looks quite promising, it has two defects which require further study. The first is an electrostatic charging and rupturing of the specimen which occurs when a part of the specimen screen is not included in the area of illumination. The second is a thermal drifting of the specimen screen caused by the high intensity and small cross section of the beam.

7. Some Critical Studies on the Contrast at the Resolving Limit of the Electron Microscope. JAMES HILLIER, *RCA Laboratories Division, Princeton, New Jersey.*—Previous work has shown that electron optical methods of enhancing contrast at the resolving limit of the electron microscope are almost completely ineffectual. Shadow-casting is limited to the study of specimen surfaces and recent observations indicate that it is unreliable for surface structures below 50A. Quantitative measurements show that the sensitivities of the photographic emulsions commonly used are too great, the number of electrons required to expose a picture element being too small to permit the detection of intensity variations of less than 50 percent. This suggests the use of finer-grained, less sensitive emulsions. Some results obtained with Eastman Kodalith Emulsion are presented and some new difficulties described.

8. Some Attempts to Produce Structureless Films, Replicas, and Shadows. S. G. ELLIS* AND JAMES HILLIER, *RCA Laboratories Division, Princeton, New Jersey.*—The interpretation of high contrast pictures of thin specimens mounted on collodion films is confused by the presence of the collodion structure. Thin evaporated films of graphite have been examined in the electron microscope, but are not completely structureless. Some attempts to produce films and replicas by polymerization induced by electron bombardment are reported, together with some observations on structure in metal shadowing.

9. The Sensitometry of Photographic Materials Exposed to Electrons. A. L. SCHOEN* AND J. F. HAMILTON, *Research Laboratories, Eastman Kodak Company, Rochester, New York.*—The wide variety of specimens examined in the electron microscope makes it desirable for the operator to know the characteristics of the photographic emulsions available. Although relative speed and contrast studies have been conducted previously on a time-scale basis, by successively uncovering additional areas on the plate to a uniformly illuminated field, this procedure is actually far from that encountered in practice where the object varies the intensity of the beam falling on the photographic plate. In the former case, the paths of almost all the electrons contributing to exposure are unobstructed and the effect of low velocity, scattered electrons is missed entirely. In order to simulate actual operating conditions, a shadow-cast metallic step wedge has been prepared, with which the emulsions may be compared. The methods of preparation are described and some of the results and possibilities of the technique are discussed.

10. Electron Microscope Pictures in Color. GEORGE W. COCHRAN,* *Utah State Agricultural College, Logan, Utah,* AND JOHN CHIDESTER, *Dugway Proving Ground, Tooele, Utah.*—The discovery by Bensley that colors could be produced by the slow development of exposed colloidal silver in photographic emulsions has been applied to the electron microscope photographic process. Extensive studies have been made to determine which conditions of exposure, original development, rehaliding, re-exposure, and second development give the most desirable colors without loss of detail or contrast. Medium contrast lantern slide plates were used for most of the color work. This color process may be used in any electron microscope laboratory without changes in instrumentation. Brilliant color transparencies have been produced which may be shown by projection or may serve as a basis for the production of dye

transfer prints. Specimen features which heretofore have been rendered only in shades of gray can now be shown in beautiful colors. There is no correlation between the actual specimen color and the colors produced by this process.

11. A Technique for Retaining Three-Dimensional Structures in Drying Specimens for the Electron Microscope.† THOMAS F. ANDERSON, *Johnson Foundation, University of Pennsylvania, Philadelphia, Pennsylvania.*—The causes of distortion in biological and colloidal specimens are discussed, together with the methods which have been used to reduce their effects. The critical point method,[1] which eliminates effects due to the surface tension of the ambient liquid, is described in detail. This method involves replacing the water of the specimen by each of a series of *miscible* liquids, the last in the series being a liquid like carbon dioxide under pressure. Then the temperature is raised above the critical point of this last liquid (31°C for CO_2), whereupon it changes imperceptibly into a gas. This gas is then slowly released to leave the specimen dry and undistorted by the passage of a phase boundary through it. Stereoscopic views of a number of such specimens will be shown to illustrate both the advantages and limitations of the method. The principal advantage is the preservation of the gross structure of the specimen. For example, ghosts of human red cells are still spherical after having been dried in this way, and corkscrew-shaped bacteria retain their corkscrew shapes. The tadpole-shaped bacteriophage T2 is seen to be adsorbed on host cells by the "tail." The "heads" of T2 and T5 bacteriophages are hexagonal in cross section when dried in this way. The gross structures of gells are preserved, too, and in stereoscopic views one may look down through the fibrous meshwork of which they are composed. The limitations of the method are (1) the possible extraction of components of the specimen by the liquid used and (2) the tendency of the specimen to be distorted in reducing its own surface area.

† This work was aided by a contract between the Office of Naval Research, Department of the Navy, and the University of Pennsylvania (NR-135-197), Physical, Chemical, and Biophysical Characterization of Viruses and Virus Systems. Reproduction in whole or in part is permitted for any purpose by the United States Government. Mr. Carl F. Oster, Jr., has rendered technical assistance which is gratefully acknowledged.

[1] T. F. Anderson, J. Appl. Phys. **21**, 724 (1950); Biol. Bull. **99**, 315–316 (1950); C. R. congrès international de microscopie électronique, Paris (1950), to be published; Trans. N. Y. Acad. Sci. (II) **13**, 130–134 (1951).

Applied Electron Microscopy

CHAIRMAN: C. J. BURTON

12. (*a*) The Examination of Fiber Replicas Using the Objective Designed by Hillier for Ultra Thin Sections and (*b*) A Focusing Microscope for the Model "B" Electron Microscope. FREDERICK F. MOREHEAD, *American Viscose Corporation, Marcus Hook, Pennsylvania.*—(*a*) The lens designed by Hillier for ultra thin sections has proved very useful for the study of fiber replicas. The inherent high contrast of this lens eliminates the necessity of metal shadowing in many cases, which is of special value in the investigation of fiber surfaces. A device for holding and orienting fibers prior to replication is described. Examples of different textile fiber replicas are given which include several types of replicating material.

(*b*) A simple and inexpensive microscope for focusing on the final viewing screen of Model "B" Electron Microscope is described. Magnification 10–15×. Most of the parts are available commercially.

13. Special Techniques of Surface Replication. C. J. CALBICK, *Bell Telephone Laboratories, Murray Hill, New Jersey.*—Evaporated film replicas can be made[1] of any suitable material, provided that the direction of incidence of the atom beam differs from that of the electron beam. Portrayal of surface topography is due in large part to *shading* caused by locally variable thickness of the replica in the direction of the electron beam. Shading must be distinguished from shadowing, which because of the complex nature of shadow boundaries is often difficult to interpret. If every atom sticks where it first strikes the surface, the shading may be called incidence-shading. Since the deposited thickness is vectorial in nature, thicknesses due to multiple sources may be vectorially added; and hence, from a shading standpoint, multiple sources or sources of extended area are equivalent to a single properly placed source. A special case of this generalization is the portrait shadow-casting recently reported.[2] Diffusion-shading, due to migration of the condensing material over the surface, occurs with silica,[3] SiO, and germanium; but a part of the material sticks where it first strikes, giving rise to incidence shading and shadows. Replication of complex surfaces of porous or reentrant nature poses a difficult problem. A ring source of silica formed by painting a 4-inch diameter tungsten loop with colloidal silica may be used to substantially eliminate shadows and provide a continuous film over the complex surface. A second evaporation of germanium, chromium, or some other heavy element at an angle of incidence between 30° and 60°, provides shading, shadowing, and high resolution of surface details. Instead of silica, the base film may be Al-Be alloy[4] evaporated from multiple sources distributed in azimuth. Stereoscopic observation is essential to interpretation of complex structure.

[1] C. J. Calbick, Bell System Tech. J. (October, 1951).
[2] D. E. Philpott, J. Appl. Phys. **22**, 982 (1951).
[3] C. J. Calbick, J. Appl. Phys. **19**, 119 (1948).
[4] W. Kaye, J. Appl. Phys. **20**, 1209 (1949).

14. The Structure of Sputtered Silver Films. C. E. ELLS AND G. D. SCOTT,* *Department of Physics, University of Toronto, Toronto, Canada.*—The structure of sputtered silver films has been compared with that of vacuum-evaporated films. For the same rate of formation the sputtered films are observed to be more continuous. When compared with evaporated silver films which are produced at much faster rates, the sputtered films exhibit a cross-over phenomenon in which the thinner films are more continuous and the thicker less continuous than evaporated films of the same thickness. Resistivity measurements confirm these effects.

15. Sputtering Etching of Metals. R. M. FISHER, *United States Steel Company Research Laboratory, Kearny, New Jersey.*—Interest has recently been revived in cathodic sputtering as a metallurgical etchant. This report describes the experimental procedures in etching samples, including innovations made at this laboratory, and some applications to ferrous metals.

16. Evaporated Replicas of Frozen Liquids at Low Temperature. H. T. MERYMAN[1] AND E. KAFIG, *Naval Medical Research Institute, National Naval Medical Center, Bethesda, Maryland.*—Silicon monoxide replicas are formed by vacuum evaporation at liquid nitrogen temperature. The apparatus permits heat etching of either surfaces or internal fracture planes. Studies have been made of ice, salt solutions, and detergents, both external and fracture surfaces. The limitations and applications of the technique will be discussed.

[1] Will not be presented; Lt. Meryman assigned to Korean duty.

17. An Empirical Survey of Inorganic Compounds Useful as Stains. A. G. RICHARDS, *University of Minnesota, St. Paul Minnesota.*—Using various types of insect cuticle as test material, a survey was made in an effort to locate favorable stains for electron microscopy. Data will be presented on the favorable results obtained with aqueous solutions (usually 2 percent) of iridium tetrachloride, zirconyl nitrate, thorium nitrate, ammoniacal silver nitrate, lead acetate, and sodium zinc uranyl acetate, anhydrous solutions of titanium tetrachloride and zirconium tetrachloride, and aqueous solutions

of the standard reagents mercuric chloride, osmic acid, phosphotungstic acid, chloroplatinic acid, and Millon's reagent. Other compounds tested as aqueous solutions without obtaining encouraging results include: NaCl, KCl, $CaCl_2$, $CrCl_3$, $MnCl_2$, $MnSO_4$, $KMnO_4$, $K_3Fe(CN)_6$, $K_4Fe(CN)_6$, $FeCl_3$, $FeSO_4$, $CoCl_2$, $Na_3Co(NO_2)_6$, $NiCl_2$, $CuSO_4$, $(NH_4)_6Mo_7O_{24}$, $AgNO_3$, $CdCl_2$, $SnCl_4$, $SbCl_3$, $KSbOC_4H_4O_6$, K_2TeO_3, KI, I+KI, KIO_4, H_5IO_6, CsCl, Cs_2SO_4, $BaCl_2$, Na_2WO_4, K_2WO_4, $KAu(CN)_2$, $HgBr_2$, HgCl, $TlNO_3$, TlI, $Pb(NO_3)_2$, $NaBiO_3$, $BiC_6H_5O_7$, $ThCl_4$, $UO_2(NO_3)_2$, and $UO_2(C_2H_3O_2)_2$.

18. Controlled Fixation with Osmium Tetroxide. R. F. BAKER AND F. W. S. MODERN,* *School of Medicine, University of Southern California, Los Angeles, California, and the Long Beach Veterans Administration Hospital, Long Beach California.*—One of the disadvantages of osmic acid as a fixative is the uneven fixation resulting from a slow rate of penetration into tissue. Outer layers become dense with an excess of osmium and inner layers remain poorly fixed. This has been overcome in the present work by a freezing technique. Small pieces of tissue are removed from the organ *in situ* and quickly immersed in liquid nitrogen. They are then removed to the stage of a freezing microtome, and slices 100–200 microns in thickness are cut. These sections are put into 2 percent osmic acid while frozen, and subsequently embedded in the usual manner. Optimum fixation time is about one hour. Advantages are uniform fixation, a great saving in time throughout the processing, and the opportunity to efficiently introduce heavy metal salts, etc.

19. Preparation of Droplet Patterns from Infective Suspensions. R. C. BACKUS, *Virus Laboratory, University of California, Berkeley, California.*—An apparatus will be described for the convenient preparation of droplet patterns from suspensions containing pathogenic or noxious materials. The device operates under reduced pressure, after an initial evacuation by aspiration, droplets being formed at the orifice of suitably positioned inlets for the admittance of liquid sample and air dispersant. Several sets of specimen grids are provided to receive droplets from a succession of sprayings. The influx of air during removal of specimen grids assures confinement of aerosol to the apparatus. The complete unit, after removal of specimen grids, may be autoclaved or otherwise decontaminated.

20. Preparation of Uniformly Dispersed Specimens of Particulate Matter for Examination with the Electron Microscope. T. D. GREEN,* J. B. BATEMAN, R. E. HARTMAN, C. A. SENSENEY, AND G. E. HESS, *Chemical Corps, Biological Laboratories, Camp Detrick, Frederick, Maryland.*—The transfer of monolayers of proteins to solid surfaces from a liquid substrate is now a well-established procedure. It has been found that if particulate matter is present in the protein solution from which the monolayer is to be formed, the surface streaming set up by the spreading of the protein over the surface of a liquid substrate causes the particles to be distributed uniformly throughout the available surface. They can then be transferred to a solid surface along with the protein monolayer. This procedure has been used in the preparation of specimens for electron microscopy by dipping collodion-coated wire mesh screens attached to glass microscope slides through the interface occupied by the monolayer. Micrographs of viruses, bacteria, and other bodies mounted in this manner will be shown. The application of the procedure to particle assay will be discussed.

21. Methods of Preparing Carbon Black Loaded Rubber Stocks for Electron Microscope Examination. MARY M. CHAPPUIS* AND M. H. POLLEY, *Research and Development Laboratory, Godfrey L. Cabot, Inc., Boston, Massachusetts.*—Two methods of preparing carbon black loaded rubber stocks for examination under the electron microscope will be described. The first method is adaptable to uncured rubber stocks and consists of stamping a small piece of stock, whose edge has been made tacky, onto a Formvar-covered glass slide. The Formvar film on which the thin layer of rubber has thus been deposited is stripped off onto water and picked up on a standard specimen grid. The second method is adaptable to vulcanized stocks. The carbon black loaded stock is embedded in butyl methacrylate monomer and after polymerization is sectioned on an International Minot Rotary microtome. The microtome is equipped with a special feed. A carefully fractured plate glass serves as the cutting edge. By these techniques excellent resolution of the carbon black particles embedded in the rubber matrix has been obtained.

22. Electron Microscope Study of Crystallized Titanium Dioxide Particles in Vitreous Enamels. WILLIAM R. LASKO, *National Lead Company, Titanium Division, Research Laboratories, Sayreville, New Jersey.*—This study is concerned with the electron microscope techniques employed in resolving the size, shape, and frequency of crystallized titanium dioxide particles precipitated by firing vitreous enamels. Two techniques were used for preparing specimens for the electron microscope. The first employed chromium-shadowed replicas of enamel button surfaces etched with hydrofluoric acid. The second employed the application of a powder technique to crystals obtained from vitreous enamel coated steel plates after removal by selective solution of the glassy matrix with hydrofluoric acid.

23. Study of Particulate Matter in Municipal Incinerator Stack Discharges. H. FROULA,* A. BUSH, AND R. BYLOFF, *Department of Engineering, University of California, Los Angeles, California.*—With the use of both light microscopy and electron microscopy, a study was made of the size distribution of particulate matter in the stack effluent of two municipal incinerators. Samples were collected by a technique employing thermal precipitation. Two interesting aspects of the results are the indications (1) that approximately 10^{15} particles are discharged to the atmosphere per ton of refuse burned and (2) that more than 50 percent of these particles are less than 0.05 micron in diameter.

24. Electron Microscope Study of Plasticized Latexes. E. B. BRADFORD, *Physical Research Laboratory, Dow Chemical Company, Midland, Michigan.*—Recent developments in the use of synthetic latexes in the paint, paper, and textile industries have focused attention upon the mechanism by which polymer particles coalesce into a coherent film. Many latexes are not film-forming unless they are modified by adding a plasticizer in order to produce useful films. Silicon monoxide replicas of the surfaces of air-dried cast films from latexes after various amounts of plasticizer have been added are shown. It can be observed that the poor films are composed of discrete individual particles, while in a continuous transparent film the particles have sintered into a polymeric mass. Replicas of surfaces of films produced from different types of synthetic latexes before and after plasticizer addition are shown. A pseudo-replica of the surface of a film from a formulated latex is shown which pictures the uniform dispersion of the pigment in the cast film.

25. The Fine Structure of Carbon Black. JAMES HILLIER, *RCA Laboratories Division, Princeton, New Jersey.*—Using carefully corrected objectives and taking precautions to minimize the accumulation of specimen contamination, it has been possible to demonstrate that carbon black particles consistently possess a fine structure which appears to lie below the 20A level. The rapid change in the appearance of this structure with instrument focus will be demonstrated.

Applied Electron Microscopy (continued)

CHAIRMAN: R. W. G. WYCKOFF

26. Molecular Weight Determinations of Polyvinylpyrrolidone Using the Electron Microscope. JOSEPH J. COMER* AND F. A. HAMM, *General Aniline and Film Corporation, Central Research Laboratory, Easton, Pennsylvania.*—It is the purpose of this paper to present a review of a molecular weight study of polyvinylpyrrolidone, a blood plasma extender, using the electron microscope. This type of determination of the molecular weight of polymers depends upon the resolution of single macromolecules. This paper describes the difficulties imposed on the electron microscopical technique by the substrate structure, structure introduced by shadowing metals, specimen contamination, and the various methods of formation of the presumed single macromolecule. Comparison of the electron microscopical molecular weight with values derived from viscosity, ultracentrifugation, and diffusiometry will be given.

27. An Electron Microscope Study of the Internal Structure of Cellulose Acetate Fibers and Films. VIRGIL PECK* AND WILBUR KAYE, *Research Laboratories, Tennessee Eastman Company, Kingsport, Tennessee.*—The internal structure of cellulose acetate fibers and films has been studied in the electron microscope by the use of etching and high resolution aluminum-beryllium alloy replicas. Surface regions of fibers and films have been brought into relief, and possible explanations for these regions are offered. Water in cellulose acetate causes pronounced effects upon the type of etched surface obtained. These effects may be explained on the basis of an insulation of the chain molecules from one another. The skin effect in highly redrafted fibers is discussed, and the wrinkling and cracking of the surface upon stretching is shown. Shadowgraphs of the edges of fibers are given to show their gross surface irregularities.

28. Criteria for the Characterization of Kaolinite, Halloysite, and a Related Mineral in Clays and Soils.† LUIS BRAMAO, *Estacao Agronomica Nacional, Sacavem, Portugal,* J. G. CADY AND S. B. HENDRICKS, *U. S. Department of Agriculture, Beltsville, Maryland,* AND MAX SWERDLOW,* *National Bureau of Standards, Washington, D. C.*—The methods of differential thermal analysis, powder x-ray diffraction, and electron microscopy provide criteria for identification of kaolinite and halloysite in clays. These are evaluated with regard to soil clays. (1) When the shape of the main endothermic peak on the D.T.A. curve of a kaolin mineral is expressed as the ratio of the tangents of the angles of slope of the two sides of the peak, the values of the slope ratios for halloysites are greater than 2.5 and for kaolinites are generally less than 1.8. Slope ratio can be used positively to identify kaolinite. (2) The presence of (*hkl*) x-ray reflections with neither *h*, *k*, nor *l* equal to zero, and (001) reflections having interplanar spacings from 7.15 to 7.20A, are also definite kaolinite criteria. (3) Electron microscopical observation of hexagonal plates of kaolinite and cylindrical forms of halloysite offer further definite criteria. Results of the three methods are entirely consistent for kaolin mineral specimens. (4) In the case of soil clays containing kaolin, both kaolinite and halloysite were often present as well as small particles of a more poorly crystallized material having certain structural properties of both kaolinite and halloysite, suggesting a separate kaolin mineral rather than an intermediate form.

† To be published in *Soil Science.*

29. Morphology of Homoionic Montmorillonites.† BENJAMIN M. SIEGEL, *Department of Engineering, Physics, Cornell University, Ithaca, New York.*—The electron microscope has been used in the study of several homoionic montmorillonites in collaboration with extensive investigations in the same clays of their gas and water adsorption, their low angle x-ray scattering at different water content, and their rheological properties. The electron micrographs of shadowed specimens sprayed from very dilute suspensions of the clays show that lithium and sodium montmorillonites deposit as highly dispersed platelets of 10A thickness. This thickness represents the unit cell dimension of this axis. The montmorillonites with other exchangeable ions such as cesium, magnesium, calcium, barium, thorium, and ferric deposit as aggregated platelets in which the individually stacked platelets of 10A thickness are clearly discernible.

† This work was supported by the Corps of Engineers, United States Army.

Electron Microscopical News from Our Friends from Abroad

FROM ENGLAND:

(a) Electron Microscopy in the Cavendish Laboratory. V. E. COSSLETT, *Ex-Secretary, British Electron Microscopy Group, Cavendish Laboratory, Department of Physics, University of Cambridge, England.* (20 min.).—Investigations are being pursued in three separate directions: in electron optics, in the construction of experimental instruments, and in the development of specimen techniques. In addition the electron microscope section carries out investigations for other departments of the university. In electron optics the main interest is in possibilities of correcting spherical aberration. Work is being done on lenses with curved axes, and also on the use of astigmatic components as suggested by Scherzer. On the instrumental side, a reflection electron microscope is under construction for direct examination of metallurgical specimens. An x-ray shadow microscope has been made, using magnetic lenses to produce a fine focus, which gives resolutions of optical order on living specimens. In the field of techniques, the use of replica metals for metallurgical work has received most attention. They have been used in a program on plastic deformation in metals, for the examination of fracture surfaces and for the standardization of etching techniques for electron microscopy. On the biological side, work is in progress on spermatozoa, spirochaetes, the effect of physical and chemical treatments on bacteria, plant and animal viruses, proteins, colloidal suspensions, muscle and nerve fibers, yeast cells. Brief reference will be made to the most interesting features in each field.

(b) The Development of Avian and Bovine Strains of Myco. Tuberculosis. E. M. BRIEGER AND MISS A. GLAUERT, *Papworth Hospital and Strangeways Laboratory, Cambridge,* AND V. E. COSSLETT* AND R. W. HORNE, *Cavendish Laboratory, Cambridge, England.* (10 min.)—Since the initial work on the Avian strain, reported to the Toronto meeting of EMSA in 1948, a more detailed investigation has been made of the development of its mycelial mode of growth. Phase contrast studies have shown how cautious must be the interpretation of electron microscopical observations of fixed and dried specimens. The bovine strain has been used for joint optical, electron microscopical and biological investigation of the reality and nature of granular forms of the organism. Granular material has been separated by centrifugation and special purification techniques, which is highly infectious on reinjection into experimental animals. The particles may, however, be miniature rods of a bacterial nature and not a different category of granules.

FROM HOLLAND:

(a) Advantages of the Use of Objective Lenses of Shortest Possible Focal Length in Electron Microscopy. A. C. VAN

DORSTEN* AND J. B. LE POOLE, *Research Laboratory, N. V. Philips Company, Eindhoven, Holland.* (10 min.).

(b) Method for the Determination of Field Shape in Magnetic Electron Lenses. A. C. VAN DORSTEN* AND J. B. LE POOLE, *Research Laboratory, N. V. Philips Company, Eindhoven, Holland.* (10 min.)

(c) The Use of Combined Observation by Electron Microscopy and Electron Diffraction in a Single Instrument. A. C. VAN DORSTEN* AND J. B. LE POOLE, *Research Laboratory, N. V. Philips Company, Eindhoven, Holland.* (10 min.)

FROM SWEDEN:

The Preparing of Ultrathin Tissue Sections for Electron Microscopy at High Resolution. FRITIOF S. SJÖSTRAND, *Department of Anatomy, Karolinska Institutet, Stockholm, Sweden.* (15 min.)—The resolution in electron micrographs of 0.1–0.2μ thick tissue sections is 100–200A and varies in different parts of the cells according to the differences in electron scattering power of the structural elements of the cells. A technique has been worked out to obtain sections only a few hundred angstrom units thick to make it possible to make better use of the resolving power of the electron microscope. This technique is based on the same principle as was published in 1943[1] and according to which the ultrathin sections are obtained by double sectioning. A new cutting device has been constructed to ascertain a high degree of reproducibility, which was lacking in the original very primitive apparatus. Ribbons of 0.5–1μ thick sections are first obtained according to the technique of Hillier and Gettner, using a modified Spencer rotary microtome. These sections are re-embedded in a thin collodion film on a plane surface. By means of a specially sharpened knife the thickness of the sections is reduced, through cutting away a part of the sections. What has been cut away is discarded and the thin parts of the sections still in the collodion surface are mounted for analysis in the electron microscope. As these parts have not been deformed or compressed, these factors do not represent any limitations regarding the minimum thickness obtainable as they do in ordinary sectioning.

[1] F. S. Sjöstrand, Nature, No. 3843, 725 (1943); Arkiv för Zoologi **35A**, 1 (1943).

Preservation of Tissue Cells for Electron Microscopy by Means of Freeze-Drying. FRITIOF S. SJÖSTRAND, *Department of Anatomy, Karolinska Institutet, Stockholm, Sweden.* (15 min.)—The double sectioning technique described in a previous paper makes it possible to obtain sufficient thin tissue sections for electron microscopy at a high resolution. To make full use of the possibilities of this technique the problem of preserving the tissue cells has to be worked on as one of the most important problems in biological electron microscopy today. Experiences using the freeze-drying technique will be discussed.

FROM GERMANY:

H. BOERSCH, *Physikalisch-Technische Bundesanstalt (20 b) Braunschweig.* (15 min.)—Electrons are diffracted by the object elastically and inelastically. The inelastic diffracted electrons reduce the contrast and in special cases the separation power of the electron microscope. The energy loss of the main part of the inelastic diffracted electrons is about 5–20 ev. Filters on the basis of the contra-field or of the negative charged electric single lens separated all electrons with an energy loss greater than 3 ev.

HELMUT RUSKA, *President of the German Society for Electron Microscopy, Institut für Mikromorphologie, Berlin-Dahlem, Germany.* (Presence is tentative.)

The Electron Microscopic Investigation of Bound Rubber. FRIEDRICH ENDTER, *also representing the German Society for Electron Microscopy.* (Presence is tentative.) (15 min.)—Rollmilled batches containing natural rubber and reinforcing fillers, such as carbon blacks and finely divided silicas, were extracted during 7 days by boiling benzol. The residues consisting of filler particles and bound rubber were applied in the usual manner to a woven wire specimen screen without the application of a carrier film. The electron microscope revealed for the system, reinforcing filler: bound rubber, a 3-dimensional lattice-like network in the spaces of which the nonbound rubber existed before the extraction.

FROM URUGUAY:

(*a*) Macromolecular Structure of the Contractile Protein of Bacterial Flagella. E. DE ROBERTIS* AND C. M. FRANCHI, *Department of Cell Ultrastructure, Instituto de Investigación de Ciencias Biológicas, Montevideo, Uruguay.* (Presence is tentative.) (10 min.)—Flagella from *Bacillus Brevis* are composite protein structures with a trypsin-resistant axial filament and a trypsin-sensitive peripheral sheath. The axial filaments of 100–120A in width are composed of two intertwined threads (or 50–70A coiled together in a *major helix* of about 410A. Each single thread is coiled into a *minor* or *molecular* spiral of about 95A. By changing the *p*H and by the use of Thiol inhibitors, changes in this macromolecular structure were observed. Addition of adenosintriphosphate results in thickening of the axial filaments and in a progressive contraction with new foldings and coils. This contraction may transform the filaments into tightly packed tufts. Treatment with mercaptide forming and oxidating thiol inhibitors interferes with the action of adenosintriphosphate. The probable mechanism underlying motility of flagella is discussed in the light of the macromolecular organization revealed with the electron microscope.

(*b*) Electron Microscope Observations on the Development of Nervous Tissue Cultured "in vitro." E. DE ROBERTIS* AND J. R. SOTELO, *Department of Cell Ultrastructure, Instituto de Investigación de Ciencias Biológicas, Montevideo, Uruguay.* (10 min.)—In cultured nervous tissue from 3- to 14-day old chick embryos, nerve cells, nerve fibers and the growing terminal end were studied under the electron microscope. The specific fibrillar substance of the neuron first appears in the cytoplasm as fine strands of a denser material. This is constituted by microvesicles of 300 to 500A in diameter or by short chains of them showing a tendency to fuse into double-edged cylindrical structures similar to the previously described *neurotubules.* Neurotubules may be isolated in the cytoplasm but they tend from bundles of 0.1μ to 1μ (neurofibrils). Nerve fibers are constituted by parallel packed neurotubules. Stages in the degeneration of the fiber and of the neurotubules have been observed. The terminal growing end contains a microvesicular material from which fibrogenesis of neurotubules takes place.

(*c*) Use of Aerosols for the Study of Macromolecules with the Electron Microscope. E. DE ROBERTIS,* C. M. FRANCHI, AND M. PODOLSKI, *Department of Cell Ultrastructure, Instituto de Investigación de Ciencias Biológicas, Montevideo, Uruguay.* (10 min.)—Examination of small isolated macromolecules is generally hampered by difficulties when trying to spread them over the supporting film, in such a way as to avoid clumping or the formation of monomolecular films. Spraying techniques have been used for this purpose but they do not prevent clustering of the small macromolecules within a *droplet* of several microns in diameter. A method has been devised by means of which the material is turned into an aerosol, charged electrically and electrostatically precipitated on the grid. When using a high enough dilution aerosol particles may contain a single macromolecule which can be deposited in isolation on the supporting film.

Biology

CHAIRMAN: T. F. ANDERSON

30. High-Resolution Observations of the Surface Contour of Particles of Tobacco Mosaic Virus. ROBLEY C. WILLIAMS, *Virus Laboratory, University of California, Berkeley, California.*—Techniques will be discussed which assist in the fairly routine attainment of electron micrographic resolutions of 30A or better. Micrographs have been obtained of uranium-shadowed particles of tobacco mosaic virus in an attempt to ascertain details of surface contour. Normally prepared particles show no periodic structure, either transverse or longitudinal. Frozen-dried preparations appear identical to ones dried in the usual way. Intense sonication produces very short rod segments which have presumably broken off the longer rods. These can frequently be photographed "on end," and many of them then exhibit an hexagonal cross section. The failure of the normal rod to appear hexagonal is discussed. It is proposed that certain anomalies result from the shadowing process and that the detailed appearance of uranium-shadowed particles should not always be taken literally. Evidence is given to show that the primary cause of shadowing anomalies is a migration of adsorbed oil molecules over the surface of the specimen.

31. An Improved Procedure for Counting Virus Particles in Plasma and Other Body Fluids. D. G. SHARP, *School of Medicine, Duke University, Durham, North Carolina.*—Virus can be observed and quantitatively estimated in animal body fluids and tissue extracts which contain large quantities of extraneous soluble material and in which there are as few as 10^8 to 10^6 particles per ml. The particles are sedimented from a known volume of the fluid onto a piece of 2 percent agar, to which they adhere while the supernatant fluid is washed away with isotonic saline or buffer solution. Collodion or formvar replicas of the agar surface show the virus particles clearly without salt or other dried soluble material. This procedure requires only 1 ml or less of fluid from the patient or test animal, and it is 100 to 1000 times more sensitive than the usual procedures, because all of the virus in the sample is put on a relatively small film area. Its diagnostic value when applied to the plasma of birds with erythro-myeloblastic leucosis has been demonstrated.

32. New Embedding Medium for Thin Sectioning. RICHARD F. BAKER* AND OLIVER WARREN, *School of Medicine, University of Southern California, Los Angeles, California.*—A combination of beeswax and a synthetic resin[1] has been found which has the advantages of simplicity and very superior cutting qualities for thin sectioning. Embedding is a two-stage process. Immersion in molten white beeswax is followed by a mixture of one part beeswax and two parts resin. Embedding is completed in two hours for 200-micron slices (Paper No. 18). If desired, extraction of the beeswax is accomplished with carbon tetrachloride, and of the resin with benzene.

[1] Harleco synthetic resin, Hartmann-Leddon Company, Philadelphia, Pennsylvania.

33. Electron Microscopy of Dental Tissues. SOL BERNICK, R. F. BAKER,* AND O. WARREN, *Schools of Dentistry and Medicine, University of Southern California, Los Angeles, California.*—Dental tissues such as dentin and enamel were studied under the electron microscope by the use of thin histological sections. Immature and mature teeth were collected and fixed with various fixatives. The teeth were then decalcified in a 10 percent nitric acid in formalin solution, dehydrated, and double-embedded in nitrocellulose and paraffin. The blocks were then sectioned at 0.05 micron, and embedding material was removed from the sections. Micrographs of enamel reveal that the enamel rods are irregular prisms cemented together by a united organic rod sheath and not by an interprismatic substance. Dentinal tubules were seen to consist of a protoplasmic process from the odontoblasts which is tightly circumscribed by the organic matrix of the dentin. No sheath of Neumann was noted.

34. The Validity of Cellular Structure as Seen with the Electron Microscope. F. B. BANG,* G. O. GEY, K. PAZEMENAS, AND P. SAPRANAUSKAS, *Departments of Medicine and Surgery, Johns Hopkins University School of Medicine and Hospital, Baltimore 5, Maryland.*—There is a natural tendency to view all new cellular structures revealed with the electron microscope as products of the drying procedure or of the high intensity electron bombardment in the microscope. In order to obtain exact information on the question of artifacts we have studied several constant cellular structures in standard cell lines by electron microscopy and compared the pictures with phase images of the living cells of the same line. The fibrous structure in a rat mesenchyme cell line No. 14 *pf* may be seen following fixation with osmium without previous washing or by washing with saline at a low temperature before fixation. Continued washing before fixation at temperatures of 39°C or above eliminates the fibers. Fixation with osmium vapors for 24 hours eliminates the fibrous structure. Phase movies of the same cell line show, even at the lower resolution, outlines of the same basic fibrous structure. The filamentous fibrils seen projecting from the edge and surface of cells may be seen in identical form in phase movies. Electron microscopy frequently shows them to be ribbed. Nuclear indentation and notching as seen in sections of tissue culture material prepared for electron microscopy is an exact reproduction of the phase pictures of living cells. Several individual cells were photographed by phase microscopy before and after osmium fixation and the individual cells transferred to electron microscope screens and pictures obtained of the identical areas. Lantern slides of these will be presented.

35. Concurrent Electron and Light Microscopic Evidence for the Existence of Mitochondria in Bacteria. STUART MUDD,* LOREN C. WINTERSCHIED, AND ERNST H. BEUTNER, *University of Pennsylvania, Philadelphia, Pennsylvania.*—Granules in mycobacteria were first described by Robert Koch and interpreted as spores. These are electron-scattering spheroidal to ellipsoidal bodies with smooth contours of minimal areas except when deformed by adjacent structures. The granules may be volatilized by intense electron bombardment, leaving torn surface membranes. Mycobacterial cells "fixed" by electron micrography[1] are partially protected from subsequent intense electron bombardment. Such bombardment of previously unirradiated cells volatilizes much of their protoplasm as well as the granules. Electron scattering granules have also been demonstrated in the cytoplasm of *Escherichia coli* and *Micrococcus cryophilus*. The granules described reduce tetrazolium, indicating the flavoprotein enzyme diaphorase, oxidize the Nadi reagent, indicating cytochrome oxidase, and give a range of colors with Janus green B characteristic of mitochondria.

[1] Hillier, Mudd, Smith, and Beutner, J. Bact. **60**, 641–654 (1950).

36. On the Identity of Isolated Chromosomes. A. R. T. DENUES, *The Sloan-Kettering Institute for Cancer Research, New York, New York.*—In the first of a series of experiments, a comparison has been made of chromatin threads from chicken erythrocytes disrupted by two radically different mechanical means. The Waring Blendor and the Logeman Hand Colloid Mill have been used. The Blendor combines great turbulence and shear; the Mill applies relatively mild, pure shear. The threads from these two procedures have been found by electron microscopy to be very similar in general forms, sizes, and fine structure. They are remarkably like the cytological

picture of mitotic chicken chromosomes. This evidence supports the identification of such chromatin threads with the chromosomes of the resting nucleus. It opposes the contrary conclusions of Lamb and others.

37. The Morphology and Development of Molluscum Contagiosum from Electronmicrographs of Thin Sections. WILLIAM G. BANFIELD,* HENRY BUNTING, MAURICE J. STRAUSS, AND JOSEPH L. MELNICK, *School of Medicine, Yale University, New Haven 11, Connecticut.*—Thin sections of molluscum contagiosum lesions from man were examined with the electron microscope. Cells of the Malpighian layer of these lesions were found which contained only a very few elementary bodies in their cytoplasm. These were considered to be early infected cells. Others were seen with a larger number of elementary bodies scattered throughout their cytoplasm, while still others were entirely filled with elementary bodies forming a mature inclusion as observable under the light microscope. The inclusion bodies were divided by septa into locules which contained the discrete, mature elementary bodies. The elementary bodies seemed to form within the material composing the matrix of the septa by a process of segmentation and condensation. The elementary bodies in the early infection forms differed somewhat from most of those in the larger inclusions which appeared identical with those obtainable from suspensions of the molluscum lesion. The elementary bodies themselves were sectioned and appeared in some instances to have a formed cortex with a less dense interior.

38. Electron Microscopic Study of Small Cytoplasmic Particles (Microsomes). DAVID B. SLAUTTERBACK, *The Sloan-Kettering Institute for Cancer Research, New York, New York.*—The fraction of the cell cytoplasm generally referred to as the microsome or small particle fraction has been isolated by differential centrifugation. In order to define better this fraction and to determine whether or not it is homogeneous, electron microscope studies have been undertaken of fractions from normal mouse liver and mouse hepatoma. An outline of the procedure for isolation and preparation of material is presented. Several slides illustrate the presence of at least two types of particles with distinct physical characteristics.

39. Fiber Formation in Trypsinogen Solutions. JEROME GROSS, *Massachusetts General Hospital, Boston, Massachusetts.*—Characteristically coiled threads about 120A in width with a pitch of about 500A were reported by this author as constituents of elastic tissue revealed by tryptic digestion.[1] Subsquently, DeRobertis and Franchi[2] claimed they were components of trypsinized bacterial flagellae. Reinvestigation of the problem has shown that the coiled threads and component filaments are present in Armour's crystalline trypsin and are seen after cleaning by centrifugation. Active trypsin from which the threads have been removed does not produce the fibrous elements from elastin. Examination of sterile crystalline trypsinogen (Worthington) failed to reveal these fibrous structures. However, after storage at +5°C for several days, sterile trypsinogen solutions buffered at *p*H 8.4 produced relatively large numbers of coiled threads and filaments morphologically similar to those described earlier. It is suggested that crystallized, sterilized trypsinogen or some component or contaminent thereof can undergo transformation to a highly characteristic fibrous form.

[1] J. Gross, J. Exptl. Medicine **89**, 699 (1949).
[2] DeRobertis and Franchi, Exptl. Cell Research **2**, 295 (1951).

40. Macromolecular Structures in a Bacterial Virus-Host Complex. C. LEVINTHAL* AND H. W. FISHER, *Laboratory of Physics, University of Michigan, Ann Arbor, Michigan.*—Using rapid decompression techniques,[1] it is possible to rupture the bacterium *Escherichia Coli B* without inactivating the virus T_2. In this way we have liberated the internal structures of the virus-host complex at well-defined times during the latent period. With this procedure no attempt is made to maintain the orientation of the structures which existed in the intact system. Electron micrographs of the liberated material yield the following information: (a) During the first half of the latent period no structures are found except those which are also in the uninfected organisms. (b) After about half of the latent period, fibrous structures are observed which are similar to those in the head of the fully formed virus. (c) When the first particles resembling virus are formed late in the latent period, they have heads which, after preparation for microscopy, are flatter than normal virus. (d) Just prior to normal lysis the cells contain many fully formed, infective, virus particles.

[1] Dean Fraser, Nature **167**, 33 (1951).

41. Electron Microscope Observations on Nucleolar Structure. E. BORYSKO* AND F. B. BANG, *Department of Medicine and Surgery, Johns Hopkins University School of Medicine, Baltimore, Maryland.*—In the course of studying sections of both normal and pathological tissues in the electron microscope, several morphological features of the nucleolus have been observed which may be of considerable interest to the cytologist. Although the majority of nucleoli appeared to be homogeneous masses, a large number exhibited hitherto unrecognized structural complexities. The most common complex form consisted of a tangled mass of fine threads, giving the nucleolus a "wormy" appearance. In other cases, the nucleolus appeared to be a mass of fine granules or a dense, vacuolated body. There was no evidence of a limiting membrane surrounding the nucleoli. Although seven normal and pathological tissues were studied, including chick chorioallantoic membrane, human liver, human skin, cultured malignant human cartilage cells, and cultured malignant rat fibroblast cells, no significant differences in nucleolar structure were observed, indicating that the complex forms occurred normally and were not a result of the pathological condition of the tissues.

42. A Study of the Fine Structure of *Paramecia Aurelia* "Killers" and "Sensitives." LEONARD HAMILTON, MARK GETTNER, AND C. C. STOCK, *Division of Experimental Chemotherapy, Sloan-Kettering Institute, New York, New York* [A combination of 2 papers (*a*) and (*b*)].—(*a*) To study the action of 2,6-diamino-purine on the kappa-particle of *Paramecium aurelia*, ultra-thin electron microscope sections were made of varieties 51–7K and 51–7SB. Pictures show fine structure in the macronucleus, mitochondria, trichocysts, food vacuoles, and in a body thought to be the kappa-particle. The effects of 2,6-diamino-purine on these structures will be demonstrated and discussed.

(*b*) In this study the cilia of *Paramecium aurelia* were examined in detail. In cross section a cilium appears to consist of nine fibrils in a loosely spaced ring with a central, apparently double, tenth fibril. The fibrils are enclosed in a sheath. In longitudinal section the cilium fibril generally appears homogeneous, but some suggestion of a helical structure has been obtained. An attempt was made to observe the neuro-motor system. A body below the membrane appears in cross section as a ring into which a straight fibril projects. A set of three parallel fibrillar structures appear on either side of the cilium just beneath the pellicle. Other fibrillar structures have been observed in the interior cytoplasm.

43. Structure of Nuclear Membrane in Sections of Central Nervous System.† J. FRANCIS HARTMANN, *Department of Anatomy, University of Minnesota, Minneapolis, Minnesota.*—Ultraviolet microspectrographic research has provided much information regarding synthesis and transport of nucleoprotein in nerve cells, but the method is unsuitable for eluci-

dating the role of the nuclear membrane in these processes. In the investigations here reported, sections from selected regions of the central nervous system have been studied in the electron microscope with special reference to the nuclear membrane. Information on the normal nerve cell in fetal and adult animals has been sought as basic to an understanding of experimental work. The nuclear membrane exhibits variations not referable to cell type, ranging from local thickenings with radially oriented structural units to a clear-cut bilaminar arrangement. These preliminary findings suggest that the nuclear membrane may play an active rather than a passive role in production or transformation of nucleoprotein. This hypothesis is currently being tested by experiments designed to produce physiological or pathological alteration of the nuclear membrane.

† This investigation was supported by a research grant from the National Institute of Mental Health, of the National Institutes of Health, Public Health Service, and by the Graduate Medical and Cancer Research Fund of the University of Minnesota.

44. Electron Microscopic and Other Studies on a Copper Cystine Complex. H. KAHLER,* B. J. LLOYD, JR., AND M. EDEN, *Federal Security Agency, U. S. Public Health Service, National Institutes of Health, National Cancer Institute, Bethesda, Maryland.*—Cystine added to Formvar coated copper screens produces clumps of fine fibers. Similar fibers were obtained when cupric salts were added to a cystine solution and placed on Formvar-coated nickel screens. The fibers form upon concentration or drying and do not exist as such in dilute solution. Their diameter is 85A. Various other amino acids and metallic salts were tried but none gave a similar fiber formation. Titration experiments indicated that $2H^+$ ions are released for each copper cystine complex formed, the ratio of copper to cystine in the complex being unity. The small units comprising the complex in dilute solution aggregate on vigorous blending into large fibers. The fibers disappear below *p*H 3, but large numbers are visible between *p*H 4 and *p*H 8.8.

Special Symposium on "Elementary Electron Optics and Factors Effecting the Electron Microscopical Image"

CHAIRMAN: F. A. HAMM, *General Aniline and Film Corporation, Easton, Pennsylvania.*

PANEL: R. G. PICARD AND JOHN REISNER, *R. C. A. Victor Division, Camden, New Jersey.*

B. M. SIEGEL, *Cornell University, Ithaca, New York.*

CECIL HALL, *Massachusetts Institute of Technology, Cambridge 39, Massachusetts.*

JAMES HILLIER AND S. G. ELLIS, *R. C. A. Research Division, Princeton, New Jersey.*

Subjects

(1) Simple Lens Theory
(2) Simple Instrument Theory
(3) Manipulation of Instrument
(4) Image Defects

(a) Misalignment
(b) High Voltage
(c) Lens Defects
(d) Stray Magnetic Fields
(e) Vibration